

156 906

STIC-Biotech/ChemLib

From: Chan, Christina
Sent: Monday, June 20, 2005 8:55 AM
To: Yu, Misook; STIC-Biotech/ChemLib
Subject: RE: Rush search request for 10/030,294

Please rush. Thanks Chris

Chris Chan

TC 1600 New Hire Training Coordinator and SPE 1644
(571)-272-0841
Remsen, 3E89

-----Original Message-----

From: Yu, Misook
Sent: Monday, June 20, 2005 5:48 AM
To: Chan, Christina
Subject: Rush search request for 10/030,294

Pls approv rush search. It is due this biweek.

Stic,
pls search SEQ ID NO:1.

Examiner Misook Yu, Ph.D.
571-272-0839 (Phone)
Art Unit 1642
REM-3A18 (Room)
REM-3C18 (Mail Box)

STAFF USE ONLY

Searcher: _____
Searcher Phone: 2-_____
Date Searcher Picked up: _____
Date Completed: _____
Searcher Prep/Rev. Time: _____
Online Time: _____

Type of Search

NA#: _____ AA#: _____
Interference: _____ SPDI: _____
S/L: _____ Oligomer: _____
Encode/Transl: _____
Structure#: _____ Text: _____
Inventor: _____ Litigation: _____

Vendors and cost where applicable

STN: _____
DIALOG: _____
QUESTEL/ORBIT: _____
LEXIS/NEXIS: _____
SEQUENCE SYSTEM: _____
WWW/Internet: _____
Other(Specify): _____

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005. Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 22, 2005, 14:15:40 ; Search time 23453 Seconds
(without alignments)
11726.935 Million cell updates/sec

Title: US-10-030-294-1
Perfect score: 5676
Sequence: 1 cagctgttcaggatgctcg.....tttatataatgtcttcogt 5676

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues
Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl: *
1: gb_ba: *
2: gb_htg: *
3: gb_in: *
4: gb_om: *
5: gb_ov: *
6: gb_pat: *
7: gb_ph: *
8: gb_pl: *
9: gb_pr: *
10: gb_ro: *
11: gb_sts: *
12: gb_sy: *
13: gb_un: *
14: gb_vi: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5676	100.0	5676	6 BD093089	BD093089 Gene enco
2	5676	100.0	5676	6 BD093091	BD093091 Gene enco
3	5676	100.0	5676	6 BD093092	BD093092 Gene enco
4	5676	100.0	5960	6 BD093090	BD093090 Gene enco
5	5673.4	100.0	13940	6 BD093100	BD093100 Gene enco
6	5635.6	99.3	162646	9 AC078809	AC078809 Homo sapi
7	669.4	11.8	160574	2 AC135489	AC135489 Rattus no
8	689.4	11.8	284005	2 AC113784	AC113784 Rattus no
9	662	11.7	204313	2 AC126304	AC126304 Rattus no
10	649	11.4	214853	10 AC113020	AC113020 Mus muscu
11	372.6	6.6	56268	2 AC091240	AC091240 Mus muscu
12	215	3.8	835	9 HSP63G01	AF124528 Homo sapi
13	212.4	3.7	159249	9 AC103882	AC103882 Homo sapi
14	205	3.6	60323	2 AC087334	AC087334 Homo sapi
15	204.4	3.6	173456	9 AL158198	AL158198 Human DNA
16	201.2	3.5	150846	2 AC036173	AC036173 Pan trogl
17	200.6	3.5	170837	2 AC135613	AC135613 Pan trogl
18	200	3.5	195808	2 AL591182	AL591182 Homo sapi
19	196.6	3.5	125377	2 AC118135	AC118135 Homo sapi

20	196.6	3.5	181442	2	AC134875	Homo sapi
21	196.6	3.5	202827	9	AC016907	AC016907 Homo sapi
22	196.2	3.5	219122	9	AC080038	AC080038 Homo sapi
23	196	3.5	148911	2	AL353142	AL353142 Homo sapi
24	196	3.5	273403	9	AC011498	AC011498 Homo sapi
25	195	3.4	47028	2	AC117488	AC117488 Homo sapi
26	194	3.4	163246	2	AC068401	AC068401 Homo sapi
27	194	3.4	174428	9	AC005821	AC005821 Homo sapi
28	193.4	3.4	118384	9	AF389851	AF389851 Homo sapi
29	193.2	3.4	196467	9	AC022392	AC022392 Homo sapi
30	193	3.4	139218	2	AC019289	AC019289 Homo sapi
31	193	3.4	151880	9	AP005716	AP005716 Homo sapi
32	193	3.4	166093	9	AC106013	AC106013 Homo sapi
33	193	3.4	175931	2	AC026618	AC026618 Homo sapi
34	193	3.4	180776	2	AC032043	AC032043 Homo sapi
35	193	3.4	183336	9	AC108688	AC108688 Homo sapi
36	193	3.4	291641	2	AC092893	AC092893 Homo sapi
37	191.8	3.4	179305	2	AC144411	AC144411 Homo sapi
38	191.2	3.4	131757	9	AC004540	AC004540 Homo sapi
39	190.6	3.4	79340	9	AL359473	AL359473 Human DNA
40	190.6	3.4	161416	2	AC026857	AC026857 Homo sapi
41	190.6	3.4	218074	9	HUAC002044	AC002044 Human Chr
42	190.4	3.4	157136	2	AC068251	AC068251 Homo sapi
43	190.4	3.4	165240	9	AP002795	AP002795 Homo sapi
44	190.4	3.4	165878	2	AC073382	AC073382 Homo sapi
45	190.4	3.4	175006	2	AC027589	AC027589 Homo sapi

ALIGNMENTS

BD093089 5676 bp DNA linear PAT 27-AUG-2002
Gene encoding promoter region of tumor suppressor gene p51.

BD093089
ACCESSION BD093089.1 GI:22638677
VERSION WO 0100818-A/1.
KEYWORDS Homo sapiens (human)

SOURCE
ORGANISM Homo sapiens

REFERENCE
1 (bases 1 to 5676)
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS
Sakai, T., Kagaya, S., Sato, T., Sukenaga, Y. and Fujii, H.

TITLE
Gene encoding promoter region of tumor suppressor gene p51

JOURNAL
Patent: WO 0100818-A 1 04-JAN-2001.

COMMENT
TOSHIYUKI SAKAI, NIPPON KAYAKU KK, SHIGEHIDE KAGAYA, TAKAMICHI SATO,
YOSHIKAZU SUKENAGA, HIDEJI FUJII

OS Homo sapiens (human)
PN WO 0100818-A/1
PD 04-JAN-2001

PF 28-JUN-2000 WO 2000JP004261
PR 29-JUN-1999 JP 99P 183195

PI TOSHIYUKI SAKAI, SHIGEHIDE KAGAYA, TAKAMICHI SATO, YOSHIKAZU SUKENAGA,
HIDEJI FUJII

PC C12N15/10.C12N5/10.C12Q1/68.A61K48/00.A61K45/00.A61P43/00, PC
A61P35/00

CC
FH Key Location/Qualifiers.

FEATURES
source
1..5676
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 100.0%; Score 5676; DB 6; Length 5676;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 5676; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 CAGCTGTTCAGGATGTCCTCGAAAGAGCCACCCACATGCTTCTGGACACTGGGTGT 60
|||||

Db 1 CAGCTGTTTCAGGATGTCGGAAGAGCCACCACATTCGCTTCCTGGACACTGGGTGT 60
QY 61 GACTTTGGAGGATACAGGTTTGCTGTTAAAGAACTGCCAACCTCTTCTGCCCAAT 120
Db 61 GACTTTGGAGGATACAGGTTTGCTGTTAAAGAACTGCCAACCTCTTCTGCCCAAT 120
QY 121 TGGCTCTGTTCCCTTGCATGCCCTCTTCTTGGGACACTCCCTTAAGGCATCTTCTTG 180
Db 121 TGGCTCTGTTCCCTTGCATGCCCTCTTCTTGGGACACTCCCTTAAGGCATCTTCTTG 180
QY 181 ACATTAACTTAACTATATAATGTTTATTTGATGAATTTTCAGTGACCTGGAAGAGATGGAG 240
Db 181 ACATTAACTTAACTATATAATGTTTATTTGATGAATTTTCAGTGACCTGGAAGAGATGGAG 240
QY 241 GTCMAATCAGAAGACACATGCTAAGTTGCAATGCACTGCTGTTTTCATTGAATTTAA 300
Db 241 GTCMAATCAGAAGACACATGCTAAGTTGCAATGCACTGCTGTTTTCATTGAATTTAA 300
QY 301 AGTCATTTCGAATACCATTCAGTTTACTTAAAGTTCTTAGGCCACGTTTACTCCTAATCGAT 360
Db 301 AGTCATTTCGAATACCATTCAGTTTACTTAAAGTTCTTAGGCCACGTTTACTCCTAATCGAT 360
QY 361 GTCAGACTGTAGCAAAATATTAGTCCAAAGTTTGAAGAGTTAGCAGGATCCTCTCCATGA 420
Db 361 GTCAGACTGTAGCAAAATATTAGTCCAAAGTTTGAAGAGTTAGCAGGATCCTCTCCATGA 420
QY 421 CAGAACTTTGGCTTCCACTTTACTTAAATAGAGATTGTGTGGTTGAGCTGCAGCTATGTA 480
Db 421 CAGAACTTTGGCTTCCACTTTACTTAAATAGAGATTGTGTGGTTGAGCTGCAGCTATGTA 480
QY 481 CAGAAAGTGTACATAAATTAATAACCAAACTCAGTCTCTTCAATTTGAGCAATAGT 540
Db 481 CAGAAAGTGTACATAAATTAATAACCAAACTCAGTCTCTTCAATTTGAGCAATAGT 540
QY 541 TGTGTAATTTACTCCACCACTCTCTCTTTGAAGTTTCTTCTGCTCTCTCTCACTATA 600
Db 541 TGTGTAATTTACTCCACCACTCTCTCTTTGAAGTTTCTTCTGCTCTCTCTCACTATA 600
QY 601 AATGCAGGATGACCTGGAAGGCTAGGACCTGAGGTTTCAGTTACCTCTGACACAAAGAA 660
Db 601 AATGCAGGATGACCTGGAAGGCTAGGACCTGAGGTTTCAGTTACCTCTGACACAAAGAA 660
QY 661 TCAGTTTCTCTGATCTCATAGTCACAGGCTGCCAGAGCTCTACGGACATGCAAGATCAT 720
Db 661 TCAGTTTCTCTGATCTCATAGTCACAGGCTGCCAGAGCTCTACGGACATGCAAGATCAT 720
QY 721 CTGCTTTAAGCCTCTTGTGTGGCATCTGTTGTTTCCACTGCCCCTGACTATGCTCT 780
Db 721 CTGCTTTAAGCCTCTTGTGTGGCATCTGTTGTTTCCACTGCCCCTGACTATGCTCT 780
QY 781 TTCTCTTGGTTTAAACAGAACCTTTTATTTCTTGTGAAATCTCTCTGCTCAGTCATGGTAGG 840
Db 781 TTCTCTTGGTTTAAACAGAACCTTTTATTTCTTGTGAAATCTCTCTGCTCAGTCATGGTAGG 840
QY 841 CCATCAGTCACATGATCAGGCTCTCTCGCCAAACATGSCATCTTTCTTTTGGGAAT 900
Db 841 CCATCAGTCACATGATCAGGCTCTCTCGCCAAACATGSCATCTTTCTTTTGGGAAT 900
QY 901 TGAATCTTAAGCTGAATAGCTGAAGTTTCAAAAAAGCTGTGTAATCTGACTTACGCCCTAC 960
Db 901 TGAATCTTAAGCTGAATAGCTGAAGTTTCAAAAAAGCTGTGTAATCTGACTTACGCCCTAC 960
QY 961 AGTGGCTTTGCAAGTGACTGTGCAATTCCTTATTTCTTAAGTCCCTGAAATTTAATTTAT 1020
Db 961 AGTGGCTTTGCAAGTGACTGTGCAATTCCTTATTTCTTAAGTCCCTGAAATTTAATTTAT 1020
QY 1021 CCTGTTTACAGCCTTTCTGAGATGTGTTTCTTCCAACTGCTCTTATAGTCTGT 1080
Db 1021 CCTGTTTACAGCCTTTCTGAGATGTGTTTCTTCCAACTGCTCTTATAGTCTGT 1080
QY 1081 GAAATTTTCATATTTCTTTTCATACATTTTCATGTTTGTGTTTGTGTTGTTGTTT 1140
Db 1081 GAAATTTTCATATTTCTTTTCATACATTTTCATGTTTGTGTTTGTGTTGTTT 1140

QY 1141 GGCTTTAGGTAGGCAGAACTCAGTTTCTGTTGTTTATATCCCAAGGAATCCTGATTGATACA 1200
Db 1141 GGCTTTAGGTAGGCAGAACTCAGTTTCTGTTGTTTATATCCCAAGGAATCCTGATTGATACA 1200
QY 1201 TCCTTCCCTTTTAAATAAAGTATCTAAGGCTCAAAGAGAGTAGGCTACCTCCCTGAGG 1260
Db 1201 TCCTTCCCTTTTAAATAAAGTATCTAAGGCTCAAAGAGAGTAGGCTACCTCCCTGAGG 1260
QY 1261 TCTGGAGTAAGTTAGTACCAGAGCTCGTACTAACCCCAAGGTTAGCCAACTGCTTTACAC 1320
Db 1261 TCTGGAGTAAGTTAGTACCAGAGCTCGTACTAACCCCAAGGTTAGCCAACTGCTTTACAC 1320
QY 1321 AACATTTGCTCTCTCTTCCAGAGTTATAGCAGTCTTGGAGAAAGAGCTACTATTTTGC 1380
Db 1321 AACATTTGCTCTCTCTTCCAGAGTTATAGCAGTCTTGGAGAAAGAGCTACTATTTTGC 1380
QY 1381 CAAAGACCTTCAGGAGCAACAAAGTCTGGGATATGTGATGATTGAATCTTAAATA 1440
Db 1381 CAAAGACCTTCAGGAGCAACAAAGTCTGGGATATGTGATGATTGAATCTTAAATA 1440
QY 1441 AGTTTGTGGACTTCTGGCCATAAATTGTATCTAAGACCAGATTTCAATCTTAAATAGCT 1500
Db 1441 AGTTTGTGGACTTCTGGCCATAAATTGTATCTAAGACCAGATTTCAATCTTAAATAGCT 1500
QY 1501 AAACAAACAAACAGAGATCCACAGGTTCCAGAGCTATATAAGAGTGAATTACTGATAC 1560
Db 1501 AAACAAACAAACAGAGATCCACAGGTTCCAGAGCTATATAAGAGTGAATTACTGATAC 1560
QY 1561 AGTTGACAAACATGATATATCTCAGAAACCATGGCATCAATGAGCAAAAAAATCCAGAC 1620
Db 1561 AGTTGACAAACATGATATATCTCAGAAACCATGGCATCAATGAGCAAAAAAATCCAGAC 1620
QY 1621 ACAGAAAGATACGTACCATATGCTGCTATTTATGATATTTCTAGCATTTGATTGTTCCAA 1680
Db 1621 ACAGAAAGATACGTACCATATGCTGCTATTTATGATATTTCTAGCATTTGATTGTTCCAA 1680
QY 1681 CATAGTAGCCAGTAGGCTCACATGGCTATTTCAATTTAGTTGATTTAAATTAAGTAAGA 1740
Db 1681 CATAGTAGCCAGTAGGCTCACATGGCTATTTCAATTTAGTTGATTTAAATTAAGTAAGA 1740
QY 1741 ATAAAAATTTAGCTCTTTCAGTAGGCTTAGCCATGTAATTTAGTATTTCTAGCATTTGATTGTTCCAA 1800
Db 1741 ATAAAAATTTAGCTCTTTCAGTAGGCTTAGCCATGTAATTTAGTATTTCTAGCATTTGATTGTTCCAA 1800
QY 1801 GGTGCAAAATAGAATTTTCTTTTATAACAGAAAGTCTTATTTGGAACAAATGTTCTAG 1860
Db 1801 GGTGCAAAATAGAATTTTCTTTTATAACAGAAAGTCTTATTTGGAACAAATGTTCTAG 1860
QY 1861 AAAATATACATATCTATATAAAACAAAGCAAGTCTAGTATTTCTTAAGCCAGGGG 1920
Db 1861 AAAATATACATATCTATATAAAACAAAGCAAGTCTAGTATTTCTTAAGCCAGGGG 1920
QY 1921 TGAGGGGAGATCGATTCGCAAGTGTGATGAGGAAAGTTTGGGTAATAGGGTTGTTGGA 1980
Db 1921 TGAGGGGAGATCGATTCGCAAGTGTGATGAGGAAAGTTTGGGTAATAGGGTTGTTGGA 1980
QY 1981 ATCTTGATTCGAGTAGAGGCTACTCGGTGTCTAATGTGTCACTCTCAGACTGAAACACT 2040
Db 1981 ATCTTGATTCGAGTAGAGGCTACTCGGTGTCTAATGTGTCACTCTCAGACTGAAACACT 2040
QY 2041 TGGAAATGGGGAATTTTCATTTGATGTAATTTACCTCATAAAGTAACTCTAAGAGGTCA 2100
Db 2041 TGGAAATGGGGAATTTTCATTTGATGTAATTTACCTCATAAAGTAACTCTAAGAGGTCA 2100
QY 2101 AGTGTTTTGTGGAATTTATTTTAACTCAGTTGCAATCTATTATGAGATGATTTTGTGA 2160
Db 2101 AGTGTTTTGTGGAATTTATTTTAACTCAGTTGCAATCTATTATGAGATGATTTTGTGA 2160
QY 2161 AATACATAAACATGTTTATTCATCCATTAGGTGCAATATTTTGTCTAGTCTCTGAAACAC 2220
Db 2161 AATACATAAACATGTTTATTCATCCATTAGGTGCAATATTTTGTCTAGTCTCTGAAACAC 2220

QY 2221 AGAGTGAATTAAGATAGCAAGCTCCCTCAAGCTGTTCACAATCCAGTACAGGAGATG 2280
 DB 2221 AGAGTGAATTAAGATAGCAAGCTCCCTCAAGCTGTTCACAATCCAGTACAGGAGATG 2280
 QY 2281 AGTCTATTCAAAATAGCTAGATCCAGGAGAAAGTTATAGTGACCTTACACAATAA 2340
 DB 2281 AGTCTATTCAAAATAGCTAGATCCAGGAGAAAGTTATAGTGACCTTACACAATAA 2340
 QY 2341 GTCCAGATATAATTAAGTACAGATAGAAAGTGGGAAAGTTCTTTATGTGGAATAA 2400
 DB 2341 GTCCAGATATAATTAAGTACAGATAGAAAGTGGGAAAGTTCTTTATGTGGAATAA 2400
 QY 2401 GAGGGAGAAATTTGGTCTTTGAAGATAGCAAGATGTGAATATGCGCAGATGGAGTTT 2460
 DB 2401 GAGGGAGAAATTTGGTCTTTGAAGATAGCAAGATGTGAATATGCGCAGATGGAGTTT 2460
 QY 2461 TAAACATCTCGTGGAGGGAGAAATATGATCCAGGCAAGAGCAACACGAGAAATA 2520
 DB 2461 TAAACATCTCGTGGAGGGAGAAATATGATCCAGGCAAGAGCAACACGAGAAATA 2520
 QY 2521 TGCAACTAGAGAAAGTGCATGAAGGGAGCAGTTGTAAATAATTTTCATGAATGTAA 2580
 DB 2521 TGCAACTAGAGAAAGTGCATGAAGGGAGCAGTTGTAAATAATTTTCATGAATGTAA 2580
 QY 2581 GTGAGAAGAAATTTGTATCATAGACACCTGAGTTTGGCAGAGTGCATGTTCTTGGCTCCTA 2640
 DB 2581 GTGAGAAGAAATTTGTATCATAGACACCTGAGTTTGGCAGAGTGCATGTTCTTGGCTCCTA 2640
 QY 2641 GAGTCAAGAAAGCAAAAGTGCCTTCTCCTACGTTATGCTCAGTGGTCCAAAGTCCAA 2700
 DB 2641 GAGTCAAGAAAGCAAAAGTGCCTTCTCCTACGTTATGCTCAGTGGTCCAAAGTCCAA 2700
 QY 2701 AACACCTTCTCCTTAAAGTACTTCTCTCCCTCCATACAAATCTAAAGTCTTCACA 2760
 DB 2701 AACACCTTCTCCTTAAAGTACTTCTCTCCCTCCATACAAATCTAAAGTCTTCACA 2760
 QY 2761 AACATCATTTAAACAGGAGAGTATGCTCAGAAAGCAATTTGCTTTCTTAGACTTCTAT 2820
 DB 2761 AACATCATTTAAACAGGAGAGTATGCTCAGAAAGCAATTTGCTTTCTTAGACTTCTAT 2820
 QY 2821 GTACGTTATTATTAACAATTTCTGCTTAAAGACTCTAAAGTCTTGGAAAGTTTCCAC 2880
 DB 2821 GTACGTTATTATTAACAATTTCTGCTTAAAGACTCTAAAGTCTTGGAAAGTTTCCAC 2880
 QY 2881 CTTGCACATCAAGATATAATTCATGATTTGTATAGTAACTTACCTTACCTTAAAGAAAT 2940
 DB 2881 CTTGCACATCAAGATATAATTCATGATTTGTATAGTAACTTACCTTAAAGAAAT 2940
 QY 2941 AAGGATCAACTATAAATAAAGAAATTAATGTTAAATTAATATGATTTGCCACTTATT 3000
 DB 2941 AAGGATCAACTATAAATAAAGAAATTAATGTTAAATTAATATGATTTGCCACTTATT 3000
 QY 3001 TTTCACTTGATCGTGTATGTTGATGCTACTGCTGTTCTGTTGAAATCTAGAGAGTTTG 3060
 DB 3001 TTTCACTTGATCGTGTATGTTGATGCTACTGCTGTTCTGTTGAAATCTAGAGAGTTTG 3060
 QY 3061 CCTCTTTTCTCGGTCACCTCGCCATTTATTTCCATATGCAATAGAGGCCAATCTT 3120
 DB 3061 CCTCTTTTCTCGGTCACCTCGCCATTTATTTCCATATGCAATAGAGGCCAATCTT 3120
 QY 3121 TTTTCATAATTAATTTAAATTTGTTGCCATTTAAATTTCTGTTCTCTTAGCTTAGTA 3180
 DB 3121 TTTTCATAATTAATTTAAATTTGTTGCCATTTAAATTTCTGTTCTCTTAGCTTAGTA 3180
 QY 3181 ACTTTAGGATTTTAAATAACAATAATTTGAATCATGACATACGTTTAAATGATTTATT 3240
 DB 3181 ACTTTAGGATTTTAAATAACAATAATTTGAATCATGACATACGTTTAAATGATTTATT 3240
 QY 3241 TAAATAGCTTAGGCTATAAACCCTTTTAAATTTTAAATAAATAGATGAGTGGTGGCT 3300
 DB 3241 TAAATAGCTTAGGCTATAAACCCTTTTAAATTTTAAATAAATAGATGAGTGGTGGCT 3300
 QY 3301 CATGCCTGTAAATCCCAACACTTTTGGGAAGCCGGTGGGAGGATAGCTTTGAGTCCAGCAG 3360

DB 3301 CATGCCTGTAAATCCCAACACTTTGGGAAGCCGGTGGGAGGATAGCTTTGAGTCCAGCAG 3360
 QY 3361 TTTTGAACACAGTCAAGGCAACACAGCAAGACCCCATATCTAAAAAACAACAAACAA 3420
 DB 3361 TTTTGAACACAGTCAAGGCAACACAGCAAGACCCCATATCTTAAAAAACAACAAACAA 3420
 QY 3421 AATTACCTGGGTATGTTGTCTCACTGTAGTCCAAAGCTACACAGAGAGCTGAGGAGAG 3480
 DB 3421 AATTACCTGGGTATGTTGTCTCACTGTAGTCCAAAGCTACACAGAGAGCTGAGGAGAG 3480
 QY 3481 AGGATCACTTGAGCCAGGAGGTTGAGGCTGAGTGCATCCATGAACCGCTGCTACACTC 3540
 DB 3481 AGGATCACTTGAGCCAGGAGGTTGAGGCTGAGTGCATCCATGAACCGCTGCTACACTC 3540
 QY 3541 AGTCTGGGTGACAGTGCAGAGAGCTGTCTCAAAAAATAAATAAATAAATAAATAAATAA 3600
 DB 3541 AGTCTGGGTGACAGTGCAGAGAGCTGTCTCAAAAAATAAATAAATAAATAAATAAATAA 3600
 QY 3601 AAAAAAACAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 3660
 DB 3601 AAAAAAACAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 3660
 QY 3661 TTTATTTGGAGTCTATATCCCTGGAGTTAAATTTAAATAATTTTAAAGAGTCTTCTCTCA 3720
 DB 3661 TTTATTTGGAGTCTATATCCCTGGAGTTAAATTTAAATAATTTTAAAGAGTCTTCTCTCA 3720
 QY 3721 TTTCTTAGAGAGTCCAAATTTGTAAATATCAGAGCTAGAGGAAACACTAGGGCTCGCCACT 3780
 DB 3721 TTTCTTAGAGAGTCCAAATTTGTAAATATCAGAGCTAGAGGAAACACTAGGGCTCGCCACT 3780
 QY 3781 CCAAGTGTGGTCCAAAGGACACAGCAGCATCAAGTAACTGGGAAACGTTGTAGAAATGACAG 3840
 DB 3781 CCAAGTGTGGTCCAAAGGACACAGCAGCATCAAGTAACTGGGAAACGTTGTAGAAATGACAG 3840
 QY 3841 AGTCTTAGGCTCACCCAGACCTACTGAACACAGAACTCTGCATTAACAAGATTTCTAGGT 3900
 DB 3841 AGTCTTAGGCTCACCCAGACCTACTGAACACAGAACTCTGCATTAACAAGATTTCTAGGT 3900
 QY 3901 GCCTCAGGCGCATTTAAACCTTGAAGCTCTGCACTAGAAATCTTCACTCCACCTTTC 3960
 DB 3901 GCCTCAGGCGCATTTAAACCTTGAAGCTCTGCACTAGAAATCTTCACTCCACCTTTC 3960
 QY 3961 ATTATAAATGGAACTCACTTGGGCTGTGGTCAAGGAAATTTGATTTATTTTAAATTTTCAGAA 4020
 DB 3961 ATTATAAATGGAACTCACTTGGGCTGTGGTCAAGGAAATTTGATTTATTTTAAATTTTCAGAA 4020
 QY 4021 CTTCTTATTTAGTCACTATATTTGCTAATAGCAGGAGAAAGCCAAACTCTTTAACT 4080
 DB 4021 CTTCTTATTTAGTCACTATATTTGCTAATAGCAGGAGAAAGCCAAACTCTTTAACT 4080
 QY 4081 GCAATTAACAATACTATAATTAATTAAGTAAAGCAATCTTCCCTTTAAGTTTACATTTTG 4140
 DB 4081 GCAATTAACAATACTATAATTAATTAAGTAAAGCAATCTTCCCTTTAAGTTTACATTTTG 4140
 QY 4141 TGAGCAAGCTGTTGATTTGGCTGGGCTCAGGCGGCTGTTTGTGAATTTTACAAAT 4200
 DB 4141 TGAGCAAGCTGTTGATTTGGCTGGGCTCAGGCGGCTGTTTGTGAATTTTACAAAT 4200
 QY 4201 CACAGATGTTAGCGCTCTCGGCTAAGTAAAGGAGAGAAATGTCAAGTTTAAATAGCT 4260
 DB 4201 CACAGATGTTAGCGCTCTCGGCTAAGTAAAGGAGAGAAATGTCAAGTTTAAATAGCT 4260
 QY 4261 TCTCCCTTCCATCTGCTGCAAGCAATAAATAATTTTATGAAACACATTTTTCAGT 4320
 DB 4261 TCTCCCTTCCATCTGCTGCAAGCAATAAATAATTTTATGAAACACATTTTTCAGT 4320
 QY 4321 TAGATTTACTTACAGGAAATGTCAAAATTTCTCTGAAGGCTTTAGATTTGTCTCAAC 4380
 DB 4321 TAGATTTACTTACAGGAAATGTCAAAATTTCTCTGAAGGCTTTAGATTTGTCTCAAC 4380
 QY 4381 TTTGACATCTACTGATGTCACCTATTTTACAGGTGTGTCTGTGACTAGGGGTGAAGGGA 4440

Db 4381 TTTGACATCTACTGATGCTACCTATTATTACAGGTGTGCTGTGACTAGGGGGTGAAGGA 4440
Qy 4441 AGATGTGAACCTCACCATGTTTAGTGACCGTTAGATACACAGAGTGGTTTTTTTTTCCCGCTG 4500
Db 4441 AGATGTGAACCTCACCATGTTTAGTGACCGTTAGATACACAGAGTGGTTTTTTTTTCCCGCTG 4500
Qy 4501 TTGGAGTCTATCTCACTAGCTCTGAAATCATATTTCAATTTCCAAATCCCAAAA 4560
Db 4501 TTGGAGTCTATCTCACTAGCTCTGAAATCATATTTCAATTTCCAAATCCCAAAA 4560
Qy 4561 ACCAGGATAAGTTTACAGCCCATATTACAGAAAGGAAATAAATTTATTTGTTGTTAGACTT 4620
Db 4561 ACCAGGATAAGTTTACAGCCCATATTACAGAAAGGAAATAAATTTATTTGTTGTTAGACTT 4620
Qy 4621 TCCTGATATTACACTGATTTGGGAAATATATGAACAATTTTATGGTTTCCCTTCGAAGTAG 4680
Db 4621 TCCTGATATTACACTGATTTGGGAAATATATGAACAATTTTATGGTTTCCCTTCGAAGTAG 4680
Qy 4681 GTCAGGTCAAAGCAAAACCAAAACAGCAAAAAGTGTAAAGATAGAGTGGAG 4740
Db 4681 GTCAGGTCAAAGCAAAACCAAAACAGCAAAAAGTGTAAAGATAGAGTGGAG 4740
Qy 4741 CCGACTGAGAGATTAAATAACTAGCAATTTTTTATTAACAGGCAATTTGAAATAATTT 4800
Db 4741 CCGACTGAGAGATTAAATAACTAGCAATTTTTTATTAACAGGCAATTTGAAATAATTT 4800
Qy 4801 GTGCACCTTCAGAAATTTCTCAATAATATATTTTCCAAATTTTAAATCTTTAAAGAAA 4860
Db 4801 GTGCACCTTCAGAAATTTCTCAATAATATATTTTCCAAATTTTAAATCTTTAAAGAAA 4860
Qy 4861 TTACTATATTATATGTAAGTACATGTCATGTTGTTGAGGTAGGATATTTAACTCAATAA 4920
Db 4861 TTACTATATTATGTAAGTACATGTCATGTTGTTGAGGTAGGATATTTAACTCAATAA 4920
Qy 4921 AGGTTATTTTCTTTTATTCGGGTGAGGCAAGCTTCTAAGGGGATGTGAAGGGATATCT 4980
Db 4921 AGGTTATTTTCTTTTATTCGGGTGAGGCAAGCTTCTAAGGGGATGTGAAGGGATATCT 4980
Qy 4981 CTTTCTCTTAGCTGAGAGGAAGTGAAGTCTTAAGTTTAAATATTAATCAAGGAATTTCCCT 5040
Db 4981 CTTTCTCTTAGCTGAGAGGAAGTGAAGTCTTAAGTTTAAATATTAATCAAGGAATTTCCCT 5040
Qy 5041 GTCTTTCTCTATTTGAGATTGTGACCAACAGCGGTGGCTGAAGGGAAAAGTGAAGGG 5100
Db 5041 GTCTTTCTCTATTTGAGATTGTGACCAACAGCGGTGGCTGAAGGGAAAAGTGAAGGG 5100
Qy 5101 CGGGGAGGGAGGAAATAGATGAAAAAACAACAAAACTTCCCTTAAGCAGCTCT 5160
Db 5101 CGGGGAGGGAGGAAATAGATGAAAAAACAACAAAACTTCCCTTAAGCAGCTCT 5160
Qy 5161 ACAAAACATTTTAGCCCCAGAAAATAGTCACAGAAAATCTCAATATCAAAACAGTATCCAGA 5220
Db 5161 ACAAAACATTTTAGCCCCAGAAAATAGTCACAGAAAATCTCAATATCAAAACAGTATCCAGA 5220
Qy 5221 TACAAGGAAGTGTATGTAGCTGAGCAGGAGTGACACTCATCAGCTCAGTTCAGTTTACA 5280
Db 5221 TACAAGGAAGTGTATGTAGCTGAGCAGGAGTGACACTCATCAGCTCAGTTCAGTTTACA 5280
Qy 5281 AAAGTCCAGGCTGTGAAATTTAACTCTGATGCCATTCATGCCAGCATCAATCAAGACA 5340
Db 5281 AAAGTCCAGGCTGTGAAATTTAACTCTGATGCCATTCATGCCAGCATCAATCAAGACA 5340
Qy 5341 GAGATCAGAAAGTTTACAGAGTCCCTCAGCTCCAAATTTGCCAAACAAGTGTGGCTACTA 5400
Db 5341 GAGATCAGAAAGTTTACAGAGTCCCTCAGCTCCAAATTTGCCAAACAAGTGTGGCTACTA 5400
Qy 5401 TACGTCAAGACTCTGAAAGCGGTGAGAGAGGGGAGAAACAACAGTATAGAGAGATGCCA 5460
Db 5401 TACGTCAAGACTCTGAAAGCGGTGAGAGAGGGGAGAAACAACAGTATAGAGAGATGCCA 5460
Qy 5461 GCTGGTAAAGATCGAGTGTATGAAAGTTTATGAAAGTTTATGATGAATCTCATTTGGCTAAA 5520
Db 5461 GCTGGTAAAGATCGAGTGTATGAAAGTTTATGAAAGTTTATGATGAATCTCATTTGGCTAAA 5520

Qy 5521 TCAAGAAACGCTCCGCTCTTTGCAATATGTATGAAGGAGAGAGTGCCTAACTTCTA 5580
Db 5521 TCAAGAAACGCTCCGCTCTTTGCAATATGTATGAAGGAGAGAGTGCCTAACTTCTA 5580
Qy 5581 TGTCTGATAGCAATTTGACCGCTATTGCTTTTAGCCTCCCGCTTTATATCTATATACAC 5640
Db 5581 TGTCTGATAGCAATTTGACCGCTATTGCTTTTAGCCTCCCGCTTTATATCTATATACAC 5640
Qy 5641 AGGTATTTGTATATTTTATATATAATTTGTTCTCCGT 5676
Db 5641 AGGTATTTGTATATTTTATATATAATTTGTTCTCCGT 5676

RESULT 2
BD093091/c
LOCUS 5676 bp DNA linear PAT 27-AUG-2002
DEFINITION Gene encoding promoter region of tumor suppressor gene p51.
ACCESSION BD093091
VERSION BD093091.1 GI:22638679
KEYWORDS WO 0100818-A/3.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 5676)
AUTHORS Sakai,T., Kagaya,S., Sato,T., Sukenaga,Y. and Fujii,H.
TITLE Gene encoding promoter region of tumor suppressor gene p51
JOURNAL Patent: WO 0100818-A 3 04-JAN-2001;
TOSHIYUKI SAKAI, NIPPON KAYAKU KK, SHIGEHIDE KAGAYA, TAKAMICHI SATO,
YOSHIKAZU SUKENAGA, HIDEJI FUJII
COMMENT OS Homo sapiens (human)
PN WO 0100818-A/3
PD 04-JAN-2001
PF 28-JUN-2000 WO 2000JP004261
PR 29-JUN-1999 JP 99P 183195
PI TOSHIYUKI SAKAI, SHIGEHIDE KAGAYA, TAKAMICHI SATO, YOSHIKAZU PI
SUKENAGA,
PI HIDEJI FUJII
PC C12N15/10, C12N5/10, C12Q1/68, A61K48/00, A61K45/00, A61P43/00, PC
A61P35/00
CC
FH Key Location/Qualifiers.
source 1. 5676
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 100.0%; Score 5676; DB 6; Length 5676;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 5676; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CAGCTGTTCCAGGATGCTCTGGAAAAGAACCCACCATTTGCTTCTGGACACTGGGTGT 60
Db 5676 CAGCTGTTCCAGGATGCTCTGGAAAAGAACCCACCATTTGCTTCTGGACACTGGGTGT 5617
Qy 61 GACTTTGGAGGGTATCAGGTTTGTCTGTTAAAGAAACTGCCAACTCTTCTGCCCAAT 120
Db 5616 GACTTTGGAGGGTATCAGGTTTGTCTGTTAAAGAAACTGCCAACTCTTCTGCCCAAT 5557
Qy 121 TGGCCTCTGTTCCCTTGCATGCCCTCTTTCCTTGGGACACTCCCTTAAGGCATCTTCTTG 180
Db 5556 TGGCCTCTGTTCCCTTGCATGCCCTCTTTCCTTGGGACACTCCCTTAAGGCATCTTCTTG 5497
Qy 181 ACATTAACTTAACTATAAATGTTTATTTGATGAATTTTCACTGACCTGGAAGAGATGGAG 240
Db 5496 ACATTAACTTAACTATAAATGTTTATTTGATGAATTTTCACTGACCTGGAAGAGATGGAG 5437
Qy 241 GTCAATCAGAGAAGACACATGGCTAAGGTTCGAATGCAATGCTGCTTTTTCATTGAATTA 300
Db 5436 GTCAATCAGAGAAGACACATGGCTAAGGTTCGAATGCAATGCTGCTTTTTCATTGAATTA 5377

Qy 301 AGTCATTCGAATACCAATTCAGTTTACCTTAAGTTCTAGCCAGCTTTACTCTCTAATCGAT 360
Db 5376 AGTCATTCGAATACCAATTCAGTTTACCTTAAGTTCTAGCCAGCTTTACTCTCTAATCGAT 5317
Qy 361 GTCAGAGCTAGCAAAATATTAGTCCAAAGTTGGAAGAGTTAGCAGATCTCTCCATGA 420
Db 5316 GTCAGAGCTAGCAAAATATTAGTCCAAAGTTGGAAGAGTTAGCAGATCTCTCCATGA 5257
Qy 421 CAGAACTTTGGCTTCCACTTTACTTAAATAGAGATTGTGTGTTGAGCTGCAGCTATGTA 480
Db 5256 CAGAACTTTGGCTTCCACTTTACTTAAATAGAGATTGTGTGTTGAGCTGCAGCTATGTA 5197
Qy 481 CAGAAAAAGTGTCAATTAATAAATCACCAAAATCTCAGTCTCTTCAATTTGAGCAATAGT 540
Db 5196 CAGAAAAAGTGTCAATTAATAAATCACCAAAATCTCAGTCTCTTCAATTTGAGCAATAGT 5137
Qy 541 TGGTGAATTTATCTCCACACCTCTCTCTCTTGAAGTTCTTTCTGCTCTCTCTCACTATA 600
Db 5136 TGGTGAATTTATCTCCACACCTCTCTCTCTTGAAGTTCTTTCTGCTCTCTCTCACTATA 5077
Qy 601 AATCGAGATGACCTGGAAGGCTAGGACCTGAGGTTCAAGTTACCTCCAGCAACAAAGGAAT 660
Db 5076 AATCGAGATGACCTGGAAGGCTAGGACCTGAGGTTCAAGTTACCTCCAGCAACAAAGGAAT 5017
Qy 661 TCAGTTTCTCTGATCTCATAGTCAAGGCTGCGAGAGCTCTACGGAACATGCAAGATCAT 720
Db 5016 TCAGTTTCTCTGATCTCATAGTCAAGGCTGCGAGAGCTCTACGGAACATGCAAGATCAT 4957
Qy 721 CTGCTTTAAGCCTCTGTGTGGGCACTGTGTGTTTTCACATGCGCTCTGTAAGTATGCTCT 780
Db 4956 CTGCTTTAAGCCTCTGTGTGGGCACTGTGTGTTTTCACATGCGCTCTGTAAGTATGCTCT 4897
Qy 781 TTCTGCTGTTAAACAGACCTTTATTTCTCTGAAAGCTCTCTGCTCAGTCATGCTAGG 840
Db 4896 TTCTGCTGTTAAACAGACCTTTATTTCTCTGAAAGCTCTCTGCTCAGTCATGCTAGG 4837
Qy 841 CCATCAGTCCCATGATCAGGCTCTCTGCGCAAAATGCGCATCTTTCTTTTGGGAATT 900
Db 4836 CCATCAGTCCCATGATCAGGCTCTCTGCGCAAAATGCGCATCTTTCTTTTGGGAATT 4777
Qy 901 TGAATCTTAAGCTGAATAGCTGAAGTTCAAAAAAGCTGTTGAATCTGACTTACGCTAC 960
Db 4776 TGAATCTTAAGCTGAATAGCTGAAGTTCAAAAAAGCTGTTGAATCTGACTTACGCTAC 4717
Qy 961 AGTGGCTTTGCAAGTGAAGTCTCATCTCTATTTCTTAAGTCCCTGAAATTAATATTTAT 1020
Db 4716 AGTGGCTTTGCAAGTGAAGTCTCATCTCTATTTCTTAAGTCCCTGAAATTAATATTTAT 4657
Qy 1021 CCTGTTTACAGCCCTTTCTGAGATGTGTGTTTCTTCCAACTGTCTCTTATAGTCTGT 1080
Db 4656 CCTGTTTACAGCCCTTTCTGAGATGTGTGTTTCTTCCAACTGTCTCTTATAGTCTGT 4597
Qy 1081 GAAATTTTCATATTTCTTTTCATACATTTTCAATGTTTGTGTTGTTGTTGTTGTTT 1140
Db 4596 GAAATTTTCATATTTCTTTTCATACATTTTCAATGTTTGTGTTGTTGTTGTTGTTT 4537
Qy 1141 GCGTTTAGGTAGGAGAAATCATGTTCTGTGTTTATACCCAAAGAAATCCCTGAATGATACA 1200
Db 4536 GCGTTTAGGTAGGAGAAATCATGTTCTGTGTTTATACCCAAAGAAATCCCTGAATGATACA 4477
Qy 1201 TCCCTTCCCTTTTAAATAAAGTATCTAAGGCTCAAAGAGAGTAGGCTACCTGCTGAGG 1260
Db 4476 TCCCTTCCCTTTTAAATAAAGTATCTAAGGCTCAAAGAGAGTAGGCTACCTGCTGAGG 4417
Qy 1261 TCTGGAGTAAAGTTAGTACAGAGCTCGTACTAACCCAGGTTAGCCAACTGCTTTACAC 1320
Db 4416 TCTGGAGTAAAGTTAGTACAGAGCTCGTACTAACCCAGGTTAGCCAACTGCTTTACAC 4357
Qy 1321 AACATTTGCTCTCTCTTTCAGAGTTATAGAGTCTTGGAGAAAGAGGCTACTATTTTTC 1380
Db 4356 AACATTTGCTCTCTCTTTCAGAGTTATAGAGTCTTGGAGAAAGAGGCTACTATTTTTC 4297

Qy 1381 CAAAGACCTCAGAGGACCAAGAAAGTTCTGGGATATGTGATGATTGAACTCTTAAAA 1440
Db 4296 CAAAGACCTCAGAGGACCAAGAAAGTTCTGGGATATGTGATGATTGAACTCTTAAAA 4237
Qy 1441 AGTTTCTGACCTCTGGCCATAATTTGTATCTAAGACCAGATTTTCAATCTTAAATAGCT 1500
Db 4236 AGTTTCTGACCTCTGGCCATAATTTGTATCTAAGACCAGATTTTCAATCTTAAATAGCT 4177
Qy 1501 AAACAAAACAAAGAGATCCACAGGTTCAAGAGCTATATAAAGAGTGAATTTACTGTATC 1560
Db 4176 AAACAAAACAAAGAGATCCACAGGTTCAAGAGCTATATAAAGAGTGAATTTACTGTATC 4117
Qy 1561 AGTTGACACATGAATATATCTCAGAAAACATGGCATCAATGAGCAAAAAAATCCAGAC 1620
Db 4116 AGTTGACACATGAATATATCTCAGAAAACATGGCATCAATGAGCAAAAAAATCCAGAC 4057
Qy 1621 ACAGAGATACCTTACCATAATGCTCCTATTTATGTGATATTTCTAGCAATTTGTTGTCGA 1680
Db 4056 ACAGAGATACCTTACCATAATGCTCCTATTTATGTGATATTTCTAGCAATTTGTTGTCGA 3997
Qy 1681 CATAGTAGCCAGTAGGCTCACATGGCTATTCAAAATTTAAGTTGATTTAAAAATTAAGTAAGA 1740
Db 3996 CATAGTAGCCAGTAGGCTCACATGGCTATTCAAAATTTAAGTTGATTTAAAAATTAAGTAAGA 3937
Qy 1741 ATAAAAATTTAGCTCTTCAAGTAGGGTACCAATGTAACTAGTGGCTACCAATCAGAC 1800
Db 3936 ATAAAAATTTAGCTCTTCAAGTAGGGTACCAATGTAACTAGTGGCTACCAATCAGAC 3877
Qy 1801 GGTGCAAAATAGAAATATTTCTTATTAACAGAAAGTTCTTATTTGGAACAAATGTTCTAG 1860
Db 3876 GGTGCAAAATAGAAATATTTCTTATTAACAGAAAGTTCTTATTTGGAACAAATGTTCTAG 3817
Qy 1861 AAAATATACATAAATCTATAAAAAACAAGCAAGTCAAGTGTCTTAAGGCCAGGG 1920
Db 3816 AAAATATACATAAATCTATAAAAAACAAGCAAGTCAAGTGTCTTAAGGCCAGGG 3757
Qy 1921 TGAGGGAGATTCGATTTGCAAAAGTGTATAGGAAAGTTTGGGGTAATAGGTTGTTGGA 1980
Db 3756 TGAGGGAGATTCGATTTGCAAAAGTGTATAGGAAAGTTTGGGGTAATAGGTTGTTGGA 3697
Qy 1981 ATCTGATTTGCGATGAAGGCTACTCGGTGTCTAATGTGTCACTCTCAGACTGAACT 2040
Db 3696 ATCTGATTTGCGATGAAGGCTACTCGGTGTCTAATGTGTCACTCTCAGACTGAACT 3637
Qy 2041 TGAATTTGGGGAATTTTCAATTTGATATAATATACCTCATAAAGTAACTCTAAGAGTCA 2100
Db 3636 TGAATTTGGGGAATTTTCAATTTGATATAATATACCTCATAAAGTAACTCTAAGAGTCA 3577
Qy 2101 AGTGTTTTGGGAAATTTTAAATCAGTTGCAATATCTTATATGAGATGATTTTGGCA 2160
Db 3576 AGTGTTTTGGGAAATTTTAAATCAGTTGCAATATCTTATATGAGATGATTTTGGCA 3517
Qy 2161 AATACATAACATGTTTATTCATCCATTTAGGTGCAATATTTTGGCTAGCTCTGAAACAC 2220
Db 3516 AATACATAACATGTTTATTCATCCATTTAGGTGCAATATTTTGGCTAGCTCTGAAACAC 3457
Qy 2221 AGAGTCAATTTAGAAATAGCAAGCTGCTCAAGCTGTTCACAAATCCAGTACAGAGATG 2280
Db 3456 AGAGTCAATTTAGAAATAGCAAGCTGCTGCTCAAGCTGTTCACAAATCCAGTACAGAGATG 3397
Qy 2281 AGTCTATTCAAAATAGCTAGACTCCAGGAGAAAGTTATAGGTGACCTTACACAAAAA 2340
Db 3396 AGTCTATTCAAAATAGCTAGACTCCAGGAGAAAGTTATAGGTGACCTTACACAAAAA 3337
Qy 2341 GTGCAGATATAATTTAGTAGCACTGAGGAGGTTTCTTTTATGTTGGAAGAA 2400
Db 3336 GTGCAGATATAATTTAGTAGCACTGAGGAGGTTTCTTTTATGTTGGAAGAA 3277
Qy 2401 GAGGAGAAATTTTGGTCTTTGAGGATGAGCAAGATGTGAATGCGCAGATGAGGTTT 2460
Db 3276 GAGGAGAAATTTTGGTCTTTGAGGATGAGCAAGATGTGAATGCGCAGATGAGGTTT 3217
Qy 2461 TAAAAATTTCTGCTGGAGGCGAGAAATATGATCCAGGCAAGAGCAACAGAAAAATA 2520

|||||
3216 TAAAACTTCTGTGGAGGCGAGAAATATGATCCAAAGCGCAAGAGCAACCGAGAAAATA 3157
|||||
QY TGCACCTAGAGGAAAGTGATGAAGGGGAGCGAGTTGTGTAAAAATAATTTTCATGAATGTAA 2580
|||||
Db TGCACCTAGAGGAAAGTGATGAAGGGGAGCGAGTTGTGTAAAAATAATTTTCATGAATGTAA 3097
|||||
QY GTGAGAGAAATTTGTATCATAGACACCTGAGTTTGGCAGAGTGCATGTTCTTGGCTCCCTA 2640
|||||
Db GTGAGAGAAATTTGTATCATAGACACCTGAGTTTGGCAGAGTGCATGTTCTTGGCTCCCTA 3037
|||||
QY GGAGTCAAGAAGAAACAAAGTGTCCCTTTCTCTACGTTATGCTCAGTGGTCCAAAGTCCAA 2700
|||||
Db GGAGTCAAGAAGAAACAAAGTGTCCCTTTCTCTACGTTATGCTCAGTGGTCCAAAGTCCAA 2977
|||||
QY AACACCTTTCTCTTAAGTACTTTTCTTCTGCCCTCATACAAAATAAAGTCTTCCACA 2760
|||||
Db AACACCTTTCTCTTAAGTACTTTTCTTCTGCCCTCATACAAAATAAAGTCTTCCACA 2917
|||||
QY AACATCAATTTAAACAGGCGAGTCAATGCTCAGAAAGCAATTTGCTTTCCTAGACTTCTAT 2820
|||||
Db AACATCAATTTAAACAGGCGAGTCAATGCTCAGAAAGCAATTTGCTTTCCTAGACTTCTAT 2857
|||||
QY GTACGTTATTATATTACAAATTTCTGCCCTAAAGACTCTAAAGTCTTGGAAAAAGTTTCCAC 2880
|||||
Db GTACGTTATTATATTACAAATTTCTGCCCTAAAGACTCTAAAGTCTTGGAAAAAGTTTCCAC 2797
|||||
QY CTTGCACATCAAGATATAATTCATGCAATTTGTATAGTAACCTTAGTCCCTCAAGAGAAT 2940
|||||
Db CTTGCACATCAAGATATAATTCATGCAATTTGTATAGTAACCTTAGTCCCTCAAGAGAAT 2737
|||||
QY AAGGATCAACTATAAAATAAAGTAATTTATGTAATTAATATGATATGATGCGCACTATT 3000
|||||
Db AAGGATCAACTATAAAATAAAGTAATTTATGTAATTAATATGATATGATGCGCACTATT 2677
|||||
QY TTTCACTTGATCGTGATGTTGTCATGCTACTGTGTGTTCTGTTGAATTCCTAGAGAGTTTG 3060
|||||
Db TTTCACTTGATCGTGATGTTGTCATGCTACTGTGTGTTCTGTTGAATTCCTAGAGAGTTTG 2617
|||||
QY CTTCTTTTCTGGGTCAACTCTCGCCATTTATTTCCATAATGCAATAGAGGCAATCTT 2557
|||||
QY TTTTCAATTAATTTTAAATTTGTTGCCATTTATTTCTGTTCCCTTTAGCTTAGTA 3120
|||||
Db TTTTCAATTAATTTTAAATTTGTTGCCATTTATTTCTGTTCCCTTTAGCTTAGTA 2497
|||||
QY ACTTTAGGATTTTAAATAACAACTATTGAAATCATGACATAGCTTTTAAATGATATATT 3240
|||||
Db ACTTTAGGATTTTAAATAACAACTATTGAAATCATGACATAGCTTTTAAATGATATATT 2437
|||||
QY TAAATACGTTAGGCTATAAACCTTTTAAATTTTAAAAAATAGATGAGTGTGGCT 3300
|||||
Db TAAATACGTTAGGCTATAAACCTTTTAAATTTTAAAAAATAGATGAGTGTGGCT 2377
|||||
QY CATGCCGTAAATCCCAACCTTTTGGGAAGCGGGTCGGGAGGATAGCTTCAGTCCAGCAG 3360
|||||
Db CATGCCGTAAATCCCAACCTTTTGGGAAGCGGGTCGGGAGGATAGCTTCAGTCCAGCAG 2317
|||||
QY TTTGAGACCACTCAGGGCAACACAGCAAGACCCCATATCTTAAAAAACAACAAACAA 3420
|||||
Db TTTGAGACCACTCAGGGCAACACAGCAAGACCCCATATCTTAAAAAACAACAAACAA 2257
|||||
QY AATTACCTGGGTATGGTTGTGCTCACCTGTAGTCCAAAGCTACACAGAAAGCTGAGGCAGA 3480
|||||
Db AATTACCTGGGTATGGTTGTGCTCACCTGTAGTCCAAAGCTACACAGAAAGCTGAGGCAGA 2197
|||||
QY AGGATCACTTGAGCCAGGAGGTTGAGGCTGCAGTGCATGATCCATGAACGCGCTGTACACTC 3540
|||||
Db AGGATCACTTGAGCCAGGAGGTTGAGGCTGCAGTGCATGATCCATGAACGCGCTGTACACTC 2137
|||||
QY AGTCTGGGTGACAGTGAAGAAGCTGTCTCAAAAATAAATAAATAAATAAATAAATAAATAA 3600
|||||

Db AGTCTGGGTGACAGTGAAGAAGCTGTCTCAAAAAATAAATAAATAAATAAATAAATAAATAA 2077
QY AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 3660
Db AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2017
QY TTAATTTGGGAGTCTATATCCCTGGAAAGTTAAATTTAAAAATAATTTTAGAAGAGTTCTTCTCA 3720
Db TTAATTTGGGAGTCTATATCCCTGGAAAGTTAAATTTAAAAATAATTTTAGAAGAGTTCTTCTCA 1957
QY TTTTCTAGAGACGTCGAAATTTGTAATAATATCAGAGCTAGAAGGAACACTAGGGCTCGCCACT 3780
Db TTTTCTAGAGACGTCGAAATTTGTAATAATATCAGAGCTAGAAGGAACACTAGGGCTCGCCACT 1897
QY CMAAGTGTGTTCCAAAGGACCGAGCAGCATCAAGTAACCTTGGGAACGTTGTAGAAATGCGAG 3840
Db CMAAGTGTGTTCCAAAGGACCGAGCAGCATCAAGTAACCTTGGGAACGTTGTAGAAATGCGAG 1837
QY AGTCCTTAGGCTCACCCAGACCTCTGAAACAGAAATCTGCAATTAACAAGATTTCTAGGT 3900
Db AGTCCTTAGGCTCACCCAGACCTCTGAAACAGAAATCTGCAATTAACAAGATTTCTAGGT 1777
QY GCCTCAAGGACATTTAAATTTGAGAAGCTCTGCACTAGAAATCTTCACTCCACCTTTC 3960
Db GCCTCAAGGACATTTAAATTTGAGAAGCTCTGCACTAGAAATCTTCACTCCACCTTTC 1717
QY AATTATAAATTTGGAATCACTTGGGCTGTGGTCAAGGAAATTTGATTATTTTAAATTTTCAAGAA 4020
Db AATTATAAATTTGGAATCACTTGGGCTGTGGTCAAGGAAATTTGATTATTTTAAATTTTCAAGAA 1657
QY CTTTCTATTATTAGTCACTATATTTTCTAATAGCAGGGAAGAAAGCAAACTCTTTAACT 4080
Db CTTTCTATTATTAGTCACTATATTTTCTAATAGTAAAGCAATCTTCCCTTTAAGTTTACATTTTG 1537
QY TGGAGCAAGCTGTTGATTTGGCTGGGCTCAGGCCGGCTGTGTGTAATTTTCAAAAT 4200
Db TGGAGCAAGCTGTTGATTTGGCTGGGCTCAGGCCGGCTGTGTGTAATTTTCAAAAT 1477
QY CACAGATGTTAGCGCTCTCGGCTAAGTAAGAAAGAGAAAGTCAAGTTTAAATAGCT 4260
Db CACAGATGTTAGCGCTCTCGGCTAAGTAAGAAAGAGAAAGTCAAGTTTAAATAGCT 1417
QY TCTCCCTTCCATCCTGGCTGAAGCAACAAATAAATAATTTTATGAAACACACATTTTGAGT 4320
Db TCTCCCTTCCATCCTGGCTGAAGCAACAAATAAATAATTTTATGAAACACACATTTTGAGT 1357
QY TAGATTTACTTACAGGAAATGTCAAATTTCTGAAAGGGCTTTAGATGTCACAAAC 4380
Db TAGATTTACTTACAGGAAATGTCAAATTTCTGAAAGGGCTTTAGATGTCACAAAC 1297
QY TTTGACATCTACTGATGTCACCTATTTCAGGTTGTCCTGCTGACCTAGGGGTGAAGGGA 4440
Db TTTGACATCTACTGATGTCACCTATTTCAGGTTGTCCTGCTGACCTAGGGGTGAAGGGA 1237
QY AGATGTGAACCTACCATGTTAGTCCGTTAGATACACAGAGTGGTTTTTTTTTCCCCCTG 4500
Db AGATGTGAACCTACCATGTTAGTCCGTTAGATACACAGAGTGGTTTTTTTTTCCCCCTG 1177
QY TTGGAGTCTATCCTAACTGAGCTTCTGAATCATATTTCAATTTCCAAATCCACAA 4560
Db TTGGAGTCTATCCTAACTGAGCTTCTGAATCATATTTCAATTTCCAAATCCACAA 1117
QY ACCAGGATAAGTTTACAGCCCATATTTAGAAAGGAAATAAATAATTTTGTGTAGACTT 4620
Db ACCAGGATAAGTTTACAGCCCATATTTAGAAAGGAAATAAATAATTTTGTGTAGACTT 1057
QY TCTGTATTTACACTGATTTGGGAATATATGAACAAATTTATATGTTTTCTTTTCGAAGTAG 4680
Db TCTGTATTTACACTGATTTGGGAATATATGAACAAATTTATATGTTTTCTTTTCGAAGTAG 997

QY	4681	GTCAAGTCAAGCAAAACCAAAACAGCAAAACCTGTAAACATATAAGATAGAGTGGAG	4740
Db	996	GTCAAGTCAAGCAAAACCAAAACAGCAAAACCTGTAAACATATAAGATAGAGTGGAG	937
QY	4741	CCGACTGAGAGATTAATAAATAAATAGATATTTTATTAACAGGCAATTTGAATAATTT	4800
Db	936	CCGACTGAGAGATTAATAAATAAATAGATATTTTATTAACAGGCAATTTGAATAATTT	877
QY	4801	GTCACTTCAGATATTTCTCAATATATATATTTTCCCAATTTTAAATATCTTTAAAGAAA	4860
Db	876	GTCACTTCAGATATTTCTCAATATATATATTTTCCCAATTTTAAATATCTTTAAAGAAA	817
QY	4861	TTACTATATATATGTAAGTACATGTGCATGTGTGTTGAGGTAGGATATTTAACTCAATAA	4920
Db	816	TTACTATATATGTAAGTACATGTGCATGTGTGAGGTAGGATATTTAACTCAATAA	757
QY	4921	AGGTTATTTTCTTTTATTCGGGTGAGCAAAAGCTTCTAAGGGGATGTGAAGGGATATCT	4980
Db	756	AGGTTATTTTCTTTTATTCGGGTGAGCAAAAGCTTCTAAGGGGATGTGAAGGGATATCT	697
QY	4981	CTTCTCTTAGCTGAGAGGAGTGAAGTCTTAAGTTTAAATATATCAAGGAATTTCCCT	5040
Db	696	CTTCTCTTAGCTGAGAGGAGTGAAGTCTTAAGTTTAAATATATCAAGGAATTTCCCT	637
QY	5041	GTCTTTCTGCTATTTGAGATTGTGACCAACAGCGGTGTGCTGAAGGGGAACTGAAGGG	5100
Db	636	GTCTTTCTGCTATTTGAGATTGTGACCAACAGCGGTGTGCTGAAGGGGAACTGAAGGG	577
QY	5101	CGGGGAGGGAGGAAATAGATGAAAAACAAAAACAAAACTTCCCTAAGCAGCTCT	5160
Db	576	CGGGGAGGGAGGAAATAGATGAAAAACAAAAACAAAACTTCCCTAAGCAGCTCT	517
QY	5161	ACAAAAATTTTAGCCCCAGAAATAGTCACAGAAATCTCAATCAACACAGTATCCAGA	5220
Db	516	ACAAAAATTTTAGCCCCAGAAATAGTCACAGAAATCTCAATCAACACAGTATCCAGA	457
QY	5221	TACAAGGAAGTGTATGTAGCTGAGAGGAGTGTGACACTCATGAGTCAAGTTCAGTTTACA	5280
Db	456	TACAAGGAAGTGTATGTAGCTGAGAGGAGTGTGACACTCATGAGTCAAGTTCAGTTTACA	397
QY	5281	AAAGTCCAGGCTGTGAAATTTAAATCTGTATGCCATTCATGCCAGCATCCAATCAGACA	5340
Db	396	AAAGTCCAGGCTGTGAAATTTAAATCTGTATGCCATTCATGCCAGCATCCAATCAGACA	337
QY	5341	GAGATCAGAGTTCAGAGATGCCCTCAGCTCCAAATTTGCCAAACAAAGTGTGGCTACTA	5400
Db	336	GAGATCAGAGTTCAGAGATGCCCTCAGCTCCAAATTTGCCAAACAAAGTGTGGCTACTA	277
QY	5401	TACGTCAAGGACTGTGAAGCGTGAAGAGGGGGAAGCAACAGTAGAGAGGATGCCCA	5460
Db	276	TACGTCAAGGACTGTGAAGCGTGAAGAGGGGGAAGCAACAGTAGAGAGGATGCCCA	217
QY	5461	GCTGTAAGGATCGAGTGTATTGCAAGTTTATGATGAATCTCATTTGGCTAAAA	5520
Db	216	GCTGTAAGGATCGAGTGTATTGCAAGTTTATGATGAATCTCATTTGGCTAAAA	157
QY	5521	TCAAGAAACGCTCCGCTCTTTTCAAAATATATGTAAGGAGAGAGTGCCTAACTCTTA	5580
Db	156	TCAAGAAACGCTCCGCTCTTTTCAAAATATATGTAAGGAGAGAGTGCCTAACTCTTA	97
QY	5581	TGCTGTAGATGATTTGACCTTATGCTTTTAGCCTCCCGGCTTATATCTATATATACAC	5640
Db	96	TGCTGTAGATGATTTGACCTTATGCTTTTAGCCTCCCGGCTTATATCTATATATACAC	37
QY	5641	AGGTTATTTGCTATATTTTATATATTTTATATTTCTCCGT	5676
Db	36	AGGTTATTTGCTATATTTTATATATTTTATATTTCTCCGT	1

RESULT 3
BD093092/c
LOCUS

DEFINITION	Gene encoding promoter region of tumor suppressor gene p51.		
ACCESSION	BD093092		
VERSION	BD093092.1	GI:22638680	
KEYWORDS	WO 0100818-A/4.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	1 (bases 1 to 5676)		
AUTHORS	Sakai, T., Kagaya, S., Sato, T., Sukenaga, Y. and Fujii, H.		
TITLE	Gene encoding promoter region of tumor suppressor gene p51		
JOURNAL	Patent: WO 0100818-A 4 04-JAN-2001;		
COMMENT	TOSHIYUKI SAKAI, NIPPON KAYAKU KK, SHIGEHIDE KAGAYA, TAKAMICHI SATO, YOSHIKAZU SUKENAGA, HIDEJI FUJII		
OS	Homo sapiens (human)		
PN	WO 0100818-A/4		
PD	04-JAN-2001		
PF	28-JUN-2000	WO 2000JP004261	
PR	29-JUN-1999	JP 99P 183195	
PI	TOSHIYUKI SAKAI, SHIGEHIDE KAGAYA, TAKAMICHI SATO, YOSHIKAZU SUKENAGA,		
PI	HIDEJI FUJII		
PC	C12N15/10, C12N5/10, C12Q1/68, A61K48/00, A61K45/00, A61P43/00, PC		
CC	A61P35/00		
FEATURES	Key Location/Qualifiers.		
source	1..5676		
ORIGIN	/organism="Homo sapiens"		
	/mol_type="genomic RNA"		
	/db_xref="taxon:9606"		
Query Match	100.0%;	Score 5676;	DB 6; Length 5676;
Best Local Similarity	100.0%;	Pred. No. 0;	
Matches 5676;	Conservative 0;	Mismatches 0;	Indels 0; Gaps 0;
QY	1	CAGCTGTTTCAGGAGTGTCTGGAAAGAAAGCCCAACCCACATTTGCTTCGGACACTGGGTGT	60
Db	5676	CAGCTGTTTCAGGAGTGTCTGGAAAGAAAGCCCAACCCACATTTGCTTCGGACACTGGGTGT	5617
QY	61	GACTTTGGAGGTATCAGGTTTGTCTGTTTAAAGAAATGCAACCTTCTTCCGCCCAAT	120
Db	5616	GACTTTGGAGGTATCAGGTTTGTCTGTTTAAAGAAATGCAACCTTCTTCCGCCCAAT	5557
QY	121	TGCGCTCTGTTCCCTTGCATGCCCTCTTCTTGGGACACTCCCTTAAGGCATCTTCTTG	180
Db	5556	TGCGCTCTGTTCCCTTGCATGCCCTCTTCTTGGGACACTCCCTTAAGGCATCTTCTTG	5497
QY	181	ACATTAACTTAACTATAAATGTTTATTTGATGAATTTTCACTGACCTGAAGAGAGATGGAG	240
Db	5496	ACATTAACTTAACTATAAATGTTTATTTGATGAATTTTCACTGACCTGAAGAGAGATGGAG	5437
QY	241	GTCAATATCAGAAGAAGCACAATGCTAAGGTTGCAATGCACTTGCTTTTTCATTGAATTA	300
Db	5436	GTCAATATCAGAAGAAGCACAATGCTAAGGTTGCAATGCACTTGCTTTTTCATTGAATTA	5377
QY	301	AGTCATTGGAATACCATTTAGTTTACTTAAAGTTTCTAGGCCACGTTTACTCTTAATCGAT	360
Db	5376	AGTCATTGGAATACCATTTAGTTTACTTAAAGTTTCTAGGCCACGTTTACTCTTAATCGAT	5317
QY	361	GTCAAGCTGTAGCAAAATATTAGTTCCTAAAGTTCGAAGAGTTAGCAGGATCTCTTCCATGA	420
Db	5316	GTCAAGCTGTAGCAAAATATTAGTTCCTAAAGTTCGAAGAGTTAGCAGGATCTCTTCCATGA	5257
QY	421	CAGAACTTTGGCTTCCACTTTTACTTAAATAGAGATTTGTGGTTGAGCTGACGATATGTA	480
Db	5256	CAGAACTTTGGCTTCCACTTTTACTTAAATAGAGATTTGTGGTTGAGCTGACGATATGTA	5197
QY	481	CAGAAAGTGTCTATACAAATTAATAAATCACCAAACTCAGTCTCTTCAATTTGAGCAATAGT	540
Db	5196	CAGAAAGTGTCTATACAAATTAATAAATCACCAAACTCAGTCTCTTCAATTTGAGCAATAGT	5137

Qy	4921	AGGTTATTTTCTTTTATTCGGGTCAGGCAAGAGCTTCTAAGGGGATGTGAAAGGGATATCT	4980
Db	756	AGGTTATTTTCTTTTATTCGGGTCAGGCAAGAGCTTCTAAGGGGATGTGAAAGGGATATCT	697
Qy	4981	CTTTCTCTTAGCTGAGAGGAAGAGTGAAGTCTTAAGTTAAATATATCAAGGAATTTCCCT	5040
Db	696	CTTTCTCTTAGCTGAGAGGAAGAGTGAAGTCTTAAGTTAAATATATCAAGGAATTTCCCT	637
Qy	5041	GTCTTTTCTATTTGAGATTGTGACCAACACAGGGCGGTGGCTGAAGGGGAAACTGAAGGG	5100
Db	636	GTCTTTTCTATTTGAGATTGTGACCAACACAGGGCGGTGGCTGAAGGGGAAACTGAAGGG	577
Qy	5101	CGGGAGGGGAGGAAATAGATGAAAAAACAACAACAACAACAACAACAACAACAACAACA	5160
Db	576	CGGGAGGGGAGGAAATAGATGAAAAAACAACAACAACAACAACAACAACAACAACAACA	517
Qy	5161	ACAAAAATTTTAGCCCCAGAAATAGTTCACAGAAATCCTCAATCAAAACAGATATCCAGA	5220
Db	516	ACAAAAATTTTAGCCCCAGAAATAGTTCACAGAAATCCTCAATCAAAACAGATATCCAGA	457
Qy	5221	TACAAGGAAGTGTATGTAGCTGAGCAGAGGTGACACTCATCAGCTCAGTTCAGTTTACA	5280
Db	456	TACAAGGAAGTGTATGTAGCTGAGCAGAGGTGACACTCATCAGCTCAGTTCAGTTTACA	397
Qy	5281	AAAGTCCAGGCTGCTGAAATTTAAACTCTGATGCCATTTCATGCCAGCATCCAATCAGACA	5340
Db	396	AAAGTCCAGGCTGCTGAAATTTAAACTCTGATGCCATTTCATGCCAGCATCCAATCAGACA	337
Qy	5341	GAGATCAGAGTTTCCAGAGATGCTTCCAGCTCCAAATTTGCCAAACAAGTGTGGCTACTA	5400
Db	336	GAGATCAGAGTTTCCAGAGATGCTTCCAGCTCCAAATTTGCCAAACAAGTGTGGCTACTA	277
Qy	5401	TACGTCAGAGACTCTGAAGCGGTGAGAGAGGGGGAAGAACAAACAGTAGAGAGTGCCTCA	5460
Db	276	TACGTCAGAGACTCTGAAGCGGTGAGAGAGGGGGAAGAACAAACAGTAGAGAGTGCCTCA	217
Qy	5461	GCTCGTGAAGATTCGAGTCTTTATGAAAGTTTTCAGTCAATTCATGAATCTTCATGGCTAAAA	5520
Db	216	GCTCGTGAAGATTCGAGTCTTTATGAAAGTTTTCAGTCAATTCATGAATCTTCATGGCTAAAA	157
Qy	5521	TCAAGAAACGCTCGGCTCTTTTGCANAATATGATGAAGGAGAGAGTGCCTTAACTTCTA	5580
Db	156	TCAAGAAACGCTCGGCTCTTTTGCANAATATGATGAAGGAGAGAGTGCCTTAACTTCTA	97
Qy	5581	TGCTGTGATGACATTTGACCTTATTCCTTTTAGCCCTCCGGCTTATATCTATATATACAC	5640
Db	96	TGCTGTGATGACATTTGACCTTATTCCTTTTAGCCCTCCGGCTTATATCTATATATACAC	37
Qy	5641	AGGTATTTGTGTATATTTATATATATTTGTTCTCCGT	5676
Db	36	AGGTATTTGTGTATATTTATATATATTTGTTCTCCGT	1
RESULT 4			
BD093090		5960 bp DNA linear	PAT 27-AUG-2002
LOCUS		Gene encoding promoter region of tumor suppressor gene p51.	
DEFINITION			
ACCESSION	BD093090		
VERSION	BD093090.1	GI:22638678	
KEYWORDS	WO 0100818-A/2.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE			
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
TITLE	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
JOURNAL	1 (bases 1 to 5960)		
	Sakai,T., Kagaya,S., Sato,T., Sukenaga,Y. and Fujii,H.		
	Gene encoding promoter region of tumor suppressor gene p51		
	Patent: WO 0100818-A 2 04-JAN-2001.		
	TOSHIYUKI SAKAI,NIPPON KAYAKU KK,SHIGEHIIDE KAGAYA,TAKAMICHI SATO,		
	YOSHIOKAZU SUKENAGA,HIDEJI FUJII		
COMMENT	OS Homo sapiens (human)		
	PN WO 0100818-A/2		
	PD 04-JAN-2001		

Db 721 CTGCTTTAAGCCTCTTGTGGTGGCATCTGTGTGTTTCCACTGCGCTGTACCTATTGCTCT 780
 Qy 781 TTCCCTTGGTTTAAAGAACCTTTATTTTCTTCTGAAACCTCTCTGCTCAGTCATGTTAGGG 840
 Db 781 TTTCTTGGTTTAAAGAACCTTTATTTTCTTCTGAAACCTCTCTGCTCAGTCATGTTAGGG 840
 Qy 841 CCATCAGTCACATGATCAGGCTCTCTCTGGCCAAACATGCGCATCTTTCTTTTGGAAAT 900
 Db 841 CCATCAGTCACATGATCAGGCTCTCTCTGGCCAAACATGCGCATCTTTCTTTTGGAAAT 900
 Qy 901 TGAATCTTAAGCTGAATAGCTGAAGTTCAAAAAGCTGTTGAATCTGACTTACGCTTAC 960
 Db 901 TGAATCTTAAGCTGAATAGCTGAAGTTCAAAAAGCTGTTGAATCTGACTTACGCTTAC 960
 Qy 961 AGTGGCTTTGCAAGTGACTGCTCAATTCCTATTCTTAAAGCTGCTGAATTTAAATTTAT 1020
 Db 961 AGTGGCTTTGCAAGTGACTGCTCAATTCCTATTCTTAAAGCTGCTGAATTTAAATTTAT 1020
 Qy 1021 CCTGGTTACAGCCCTTTCTGAGATGTGTGTTTCTTTTCCAACTGCTCTTTATAGTCTGT 1080
 Db 1021 CCTGGTTACAGCCCTTTCTGAGATGTGTGTTTCTTTTCCAACTGCTCTTTATAGTCTGT 1080
 Qy 1081 GAATTTTCATATTTCTTTTCTATCATATTTCAATGTTTGTGTTTGTGTTGTTGTTT 1140
 Db 1081 GAATTTTCATATTTCTTTTCTATCATATTTCAATGTTTGTGTTTGTGTTGTTGTTT 1140
 Qy 1141 GGCTTTAGGTAGGAGATCAGTTCTGTTGTTTATACCCAGGAATCCTGATTGATACA 1200
 Db 1141 GGCTTTAGGTAGGAGATCAGTTCTGTTGTTTATACCCAGGAATCCTGATTGATACA 1200
 Qy 1201 TCCTTCCCTTTTAAAAATAAAGTATCTAAGCTCAAGAGAGTAGGCTACCTGCTGAGG 1260
 Db 1201 TCCTTCCCTTTTAAAAATAAAGTATCTAAGCTCAAGAGAGTAGGCTACCTGCTGAGG 1260
 Qy 1261 TCTGGAGTAAGTTPAGTACCAGAGCTCGTACTAAACCCAGGTTAGCCAACTGCTTTAC 1320
 Db 1261 TCTGGAGTAAGTTPAGTACCAGAGCTCGTACTAAACCCAGGTTAGCCAACTGCTTTAC 1320
 Qy 1321 AACATTTGCTCTCTCTCAGAGTTATAGCAGTCTTGGAGTATGATGATGAACTCTTAAAA 1380
 Db 1321 AACATTTGCTCTCTCTCAGAGTTATAGCAGTCTTGGAGTATGATGATGAACTCTTAAAA 1380
 Qy 1381 CAAAGACCTCAGGAGGACCAAGAACAAAGTCTCTGGATATGATGATGAACTCTTAAAA 1440
 Db 1381 CAAAGACCTCAGGAGGACCAAGAACAAAGTCTCTGGATATGATGATGAACTCTTAAAA 1440
 Qy 1441 AGTTTGTGGACTCTCTGGCCATATTTGTGATCTAAGACCAAGATTTCAATCTTAAATAGCT 1500
 Db 1441 AGTTTGTGGACTCTCTGGCCATATTTGTGATCTAAGACCAAGATTTCAATCTTAAATAGCT 1500
 Qy 1501 AAACAAACAAACAGAGATCCAGGTTTACGAGCTATATAGAGTGAATTTACTGTATAC 1560
 Db 1501 AAACAAACAAACAGAGATCCAGGTTTACGAGCTATATAGAGTGAATTTACTGTATAC 1560
 Qy 1561 AGTTGCAACATGATATATCTCAGAACCAATGCTATCAGCAACCAATGCAACCAATGCAAC 1620
 Db 1561 AGTTGCAACATGATATATCTCAGAACCAATGCTATCAGCAACCAATGCAACCAATGCAAC 1620
 Qy 1621 ACAGAAATACGTAACATATGCTGCAATTTATGTGATTTCTAGCATTTGTTATGTCCTCAA 1680
 Db 1621 ACAGAAATACGTAACATATGCTGCAATTTATGTGATTTCTAGCATTTGTTATGTCCTCAA 1680
 Qy 1681 CATAGTACCGAGTACGCTCAGATGCTATTTCAAAATTTAAGTTGATTTAAATTAAGTAAGA 1740
 Db 1681 CATAGTACCGAGTACGCTCAGATGCTATTTCAAAATTTAAGTTGATTTAAATTAAGTAAGA 1740
 Qy 1741 ATAAAAATTTAGCTCTTCTAGTAGGTTAGCCACATGTAAGTGGCTTACCACATCAGAC 1800
 Db 1741 ATAAAAATTTAGCTCTTCTAGTAGGTTAGCCACATGTAAGTGGCTTACCACATCAGAC 1800
 Qy 1801 GGTGCAATATAGATATTTCTTTTATACAGAAAGTTCTTATGGAACCAATGTTCTAG 1860
 Db 1801 GGTGCAATATAGATATTTCTTTTATACAGAAAGTTCTTATGGAACCAATGTTCTAG 1860

Qy 1861 AAAATATACATAATCTATATAAAAAACAAAGCAAGTCAAGTATGTTCTTAAGGCCAGGGG 1920
 Db 1861 AAAATATACATAATCTATATAAAAAACAAAGCAAGTCAAGTATGTTCTTAAGGCCAGGGG 1920
 Qy 1921 TGAGGGAGATCGATGTCGAAAGTGGTATGAGGAAAGTTTGGGTAAATAGGGTTGTTGGA 1980
 Db 1921 TGAGGGAGATCGATGTCGAAAGTGGTATGAGGAAAGTTTGGGTAAATAGGGTTGTTGGA 1980
 Qy 1981 ATCTTGATTCGATGAAGCTACTCGGTGTCTAATGTGTCACTCTCTCAGACTGAACACT 2040
 Db 1981 ATCTTGATTCGATGAAGCTACTCGGTGTCTAATGTGTCACTCTCTCAGACTGAACACT 2040
 Qy 2041 TGSAAATGCGAATTTTCTTGTATGTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2100
 Db 2041 TGSAAATGCGAATTTTCTTGTATGTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2100
 Qy 2101 AGTGTGTTGTTGGAATTTATTTTAACTCAGTTGCAATCTTATTTATGATGATGTTTGA 2160
 Db 2101 AGTGTGTTGTTGGAATTTATTTTAACTCAGTTGCAATCTTATTTATGATGATGTTTGA 2160
 Qy 2161 AATACATAACATGTTTATTTTCACTCATTAGTGTCAATTTTCTAGTCTCTGAAACAC 2220
 Db 2161 AATACATAACATGTTTATTTTCACTCATTAGTGTCAATTTTCTAGTCTCTGAAACAC 2220
 Qy 2221 AGAGATGAATTTAGATAGCAAGCTGCTCCTCAAGCTGTTTCACAATCCAGTACAGGAGATG 2280
 Db 2221 AGAGATGAATTTAGATAGCAAGCTGCTCCTCAAGCTGTTTCACAATCCAGTACAGGAGATG 2280
 Qy 2281 AGTCTATTTCAAAATAGCTAGTCTCAGGAGAAAGTTATAGGTGACCTTTACACAAAAA 2340
 Db 2281 AGTCTATTTCAAAATAGCTAGTCTCAGGAGAAAGTTATAGGTGACCTTTACACAAAAA 2340
 Qy 2341 GTCCAGATATTTTGTGTTTGAAGGATGAGCAAGATGTAATATGCGAGATGAGGTTT 2400
 Db 2341 GTCCAGATATTTTGTGTTTGAAGGATGAGCAAGATGTAATATGCGAGATGAGGTTT 2400
 Qy 2401 GAGGAGAAATTTTGTGTTTGAAGGATGAGCAAGATGTAATATGCGAGATGAGGTTT 2460
 Db 2401 GAGGAGAAATTTTGTGTTTGAAGGATGAGCAAGATGTAATATGCGAGATGAGGTTT 2460
 Qy 2461 TAAAAATCTCTGCTGGAGGAGAAATATGATCCAGGCAAGAGCAACAGCAAAAAATA 2520
 Db 2461 TAAAAATCTCTGCTGGAGGAGAAATATGATCCAGGCAAGAGCAACAGCAAAAAATA 2520
 Qy 2521 TGCAACCTAGAGAAAGTGCATGAGGGAGCAGTTGTAATAATTTTTCATGAATGTA 2580
 Db 2521 TGCAACCTAGAGAAAGTGCATGAGGGAGCAGTTGTAATAATTTTTCATGAATGTA 2580
 Qy 2581 GTGAGAAATTTTGTATCATAGACACCTGAGTTTGGCAGAGTGCATGTTCTTGGCTCTTA 2640
 Db 2581 GTGAGAAATTTTGTATCATAGACACCTGAGTTTGGCAGAGTGCATGTTCTTGGCTCTTA 2640
 Qy 2641 GGAGTCAAGAGAAACAAAGTGTCCCTTTCTCTAAGTTATGCTCAGTGTGCTCAAGTCAA 2700
 Db 2641 GGAGTCAAGAGAAACAAAGTGTCCCTTTCTCTAAGTTATGCTCAGTGTGCTCAAGTCAA 2700
 Qy 2701 AACACCTTTCTCTTAAAGTACTTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2760
 Db 2701 AACACCTTTCTCTTAAAGTACTTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2760
 Qy 2761 AACATCATTTTAAACAGGAGTGCATGCTCAGAAAGGCAATTTGCTTTCTCTAGACTTCTAT 2820
 Db 2761 AACATCATTTTAAACAGGAGTGCATGCTCAGAAAGGCAATTTGCTTTCTCTAGACTTCTAT 2820
 Qy 2821 GTAGCTTTATTTATTAATTTTCTGCTTAAAGACTCTAAAGCTCTGAAAGGTTTCCAC 2880
 Db 2821 GTAGCTTTATTTATTAATTTTCTGCTTAAAGACTCTAAAGCTCTGAAAGGTTTCCAC 2880
 Qy 2881 CTTGCAATCAAGATATTAATTTGATGATTTGATGATTTGATGATTTGATGATTTGATGAT 2940
 Db 2881 CTTGCAATCAAGATATTAATTTGATGATTTGATGATTTGATGATTTGATGATTTGATGAT 2940

QY	2941	AAGATGAACTATAAATATAAGAGTAATTAATGTTAAATTAATAATGATTGCCACTTATT	3000
Db	2941		
QY	2941	AAGATGAACTATAAATATAAGAGTAATTAATGTTAAATTAATAATGATTGCCACTTATT	3000
Db	2941		
QY	3001	TTTCACTTGATCGTGATGTTGCTACTGCTACTGGTGTCTGTTCAATTTCTAGAGAGTTTG	3060
Db	3001		
QY	3001	TTTCACTTGATCGTGATGTTGCTACTGCTACTGGTGTCTGTTCAATTTCTAGAGAGTTTG	3060
Db	3001		
QY	3061	CTCTCTTTCTCGGCTCAACTCTCGCCATTTATTTCCATATGCAATAGAGGCAATCTTT	3120
Db	3061		
QY	3121	TTTCAATAATTTATTTAAATTTGTTGCCATTTATTTCTGTTCTCTCTTAGCTTAGTA	3180
Db	3121		
QY	3121	TTTCAATAATTTATTTAAATTTGTTGCCATTTATTTCTGTTCTCTCTTAGCTTAGTA	3180
Db	3121		
QY	3181	ACTTTAGGATTTTAAATAACAACACTATTTGAAATCATGACATACGTTTAAATGATATTT	3240
Db	3181		
QY	3241	TAAATACGTTAGGCTATAAACCCTTTTAAATTTTAAAAAATAGATGATGTTGGCT	3300
Db	3241		
QY	3241	TAAATACGTTAGGCTATAAACCCTTTTAAATTTTAAAAAATAGATGATGTTGGCT	3300
Db	3241		
QY	3301	CATGCCCTGTAATCCCAACACTTTGGGAAGCGGTCGGGAGGATAGCTTCAGTCCAGCAG	3360
Db	3301		
QY	3301	CATGCCCTGTAATCCCAACACTTTGGGAAGCGGTCGGGAGGATAGCTTCAGTCCAGCAG	3360
Db	3301		
QY	3361	TTTGAGACCACTGAGGCAACACAGCAAGACCCCATATCTTAAAAAACAACAAACAA	3420
Db	3361		
QY	3421	AATTTACCTGGGTATGGTGTGCTCACCTGTAGTCCAAAGCTACACAGGAAGCTAGGCAGA	3480
Db	3421		
QY	3421	AATTTACCTGGGTATGGTGTGCTCACCTGTAGTCCAAAGCTACACAGGAAGCTAGGCAGA	3480
Db	3421		
QY	3481	AGGATCACTTTGAGCCAGGAGGTTGAGGCTGCAGTGATCCATGAACGCGCTGTACACTC	3540
Db	3481		
QY	3481	AGGATCACTTTGAGCCAGGAGGTTGAGGCTGCAGTGATCCATGAACGCGCTGTACACTC	3540
Db	3481		
QY	3541	AGTCTGGGTGACAGTGAAGAAGCTGTCTCAAAAATAATAAATAAATAAATAAATACTTTT	3600
Db	3541		
QY	3541	AGTCTGGGTGACAGTGAAGAAGCTGTCTCAAAAATAATAAATAAATAAATAAATACTTTT	3600
Db	3541		
QY	3601	AAAAACAATAATTAATTTTAAAAACACACACACTAGAGATGTTTGCAATTTGA	3660
Db	3601		
QY	3601	AAAAACAATAATTTTAAAAACACACACACTAGAGATGTTTGCAATTTGA	3660
Db	3601		
QY	3661	TTATTTGGGAGTCTATATCCCTGGAAGTTAAATTTAAAAATAATTTAGAAAGAGTCTTCTCTCA	3720
Db	3661		
QY	3661	TTATTTGGGAGTCTATATCCCTGGAAGTTAAATTTAAAAATAATTTAGAAAGAGTCTTCTCTCA	3720
Db	3661		
QY	3721	TTTCTTAGAGACGTCGAATTTGTAATAATCAGAGCTAGAAGGAAACACTTAGGGCTCGCCACT	3780
Db	3721		
QY	3721	TTTCTTAGAGACGTCGAATTTGTAATAATCAGAGCTAGAAGGAAACACTTAGGGCTCGCCACT	3780
Db	3721		
QY	3781	CCAAAGTGTGGTCCAAGGACAGCAGATCAAGTAACCTGGGAACGTTGTAGAAATGTCAG	3840
Db	3781		
QY	3781	CCAAAGTGTGGTCCAAGGACAGCAGATCAAGTAACCTGGGAACGTTGTAGAAATGTCAG	3840
Db	3781		
QY	3841	AGTCTTAGGCTCACCCAGACCTACTGAAACCCAGAACTCTCATTTAAACAAGATTTCTAGGT	3900
Db	3841		
QY	3841	AGTCTTAGGCTCACCCAGACCTACTGAAACCCAGAACTCTCATTTAAACAAGATTTCTAGGT	3900
Db	3841		
QY	3901	GCCTCACGGGCACATTTAAACTTTGAGAAGCTGTGCACCTAGAGAAATCTTCACTCCACCTTTC	3960
Db	3901		
QY	3901	GCCTCACGGGCACATTTAAACTTTGAGAAGCTGTGCACCTAGAGAAATCTTCACTCCACCTTTC	3960
Db	3901		
QY	3961	ATTATAAATGGAATCACTTGGGCTGTGGTCACAGGAAATTTGATTTTAAATTTTCAGAA	4020
Db	3961		
QY	3961	ATTATAAATGGAATCACTTGGGCTGTGGTCACAGGAAATTTGATTTTAAATTTTCAGAA	4020
Db	3961		
QY	4021	CTTCTTATTTAGGTCATCTATTTTGTCTAATAGCAGGGAAGAAAGCCAAACTCTTTAACT	4080
Db	4021		

Db	4021	CCTTCTATTTAGGTCATCTATATTTCTTAATAGCAGGGAAGAAAGCCAAACTCTTTAACT	4080
QY	4081	GCAATTAACAAATCTATAAATTAATTAAGTAAGCAATCTTCCCTTTAAAGTTTACATTTTG	4140
Db	4081		
QY	4081	GCAATTAACAAATCTATAAATTAATTAAGTAAGCAATCTTCCCTTTAAAGTTTACATTTTG	4140
Db	4081		
QY	4141	TGGAGCAAGCTGTTTGAATTTGGCTGGGCTCAGGCCGGCTGTTTCTGTAATTTTCACAAT	4200
Db	4141		
QY	4141	TGGAGCAAGCTGTTTGAATTTGGCTGGGCTCAGGCCGGCTGTTTCTGTAATTTTCACAAT	4200
Db	4141		
QY	4201	CACAGATGTTTAGCCGCTCTCGGGCTAAAGTAAGAAAGAAAGAGAAATCTCAAGTTTAAATAGCT	4260
Db	4201		
QY	4201	CACAGATGTTTAGCCGCTCTCGGGCTAAAGTAAGAAAGAAAGAGAAATCTCAAGTTTAAATAGCT	4260
Db	4201		
QY	4261	TCTCCCTTCCATCTCGGCTGAAGCAACAAATAAATAATTTTATGAAACACATTTTGAGT	4320
Db	4261		
QY	4261	TCTCCCTTCCATCTCGGCTGAAGCAACAAATAAATAATTTTATGAAACACATTTTGAGT	4320
Db	4261		
QY	4321	TAGATTTTACTTACAGGGAATGTCAAAATTTCTCTGAAGGGCTTTAGATTTGTCTCAAC	4380
Db	4321		
QY	4321	TAGATTTTACTTACAGGGAATGTCAAAATTTCTCTGAAGGGCTTTAGATTTGTCTCAAC	4380
Db	4321		
QY	4381	TTTGACATCTACTGTGTGTCACCTATTTACAGGTGTGTCTCTGACTAGGGGGTGAAGGGA	4440
Db	4381		
QY	4381	TTTGACATCTACTGTGTGTCACCTATTTACAGGTGTGTCTCTGACTAGGGGGTGAAGGGA	4440
Db	4381		
QY	4441	AGATGTGAATCTACCAATGTTAGTACCGTTAGATACACAGAGTGGTTTTTTTCCCGCTG	4500
Db	4441		
QY	4441	AGATGTGAATCTACCAATGTTAGTACCGTTAGATACACAGAGTGGTTTTTTTCCCGCTG	4500
Db	4441		
QY	4501	TTGGAGCTCTATCTTAAGCTGCTTCTGAATCATATTTTCAATTTTCCAAATCCCAAA	4560
Db	4501		
QY	4501	TTGGAGCTCTATCTTAAGCTGCTTCTGAATCATATTTTCAATTTTCCAAATCCCAAA	4560
Db	4501		
QY	4561	ACCAGGATTAAGTTTACAGCCCATATTTACAGAAAGAAATAAATTTATTTGTGTAGACTT	4620
Db	4561		
QY	4561	ACCAGGATTAAGTTTACAGCCCATATTTACAGAAAGAAATAAATTTATTTGTGTAGACTT	4620
Db	4561		
QY	4621	TCTGTATATTAACACTTTGGGAATATATGAACAAATTTTATTTGTTTCTTTCGAAAGTAG	4680
Db	4621		
QY	4621	TCTGTATATTAACACTTTGGGAATATATGAACAAATTTTATTTGTTTCTTTCGAAAGTAG	4680
Db	4621		
QY	4681	GTCAAGTCAAGCAACAAACCAAAACAGCAAACTGTAAGACATAAAGAAATACAGTGGAG	4740
Db	4681		
QY	4681	GTCAAGTCAAGCAACAAACCAAAACAGCAAACTGTAAGACATAAAGAAATACAGTGGAG	4740
Db	4681		
QY	4741	CCGACTGAGAGATTAATAATAAATACTAGAAATTTTATTTTAAACAGGCAATTTGAAATATTT	4800
Db	4741		
QY	4741	CCGACTGAGAGATTAATAATAAATACTAGAAATTTTATTTTAAACAGGCAATTTGAAATATTT	4800
Db	4741		
QY	4801	GTGCACTTCGAAATATTTTCAATAATATATTTTCCAAATTTTAAATCTTTTAAAGAAA	4860
Db	4801		
QY	4801	GTGCACTTCGAAATATTTTCAATAATATATTTTCCAAATTTTAAATCTTTTAAAGAAA	4860
Db	4801		
QY	4861	TTACTATTTATATGTAAGTACATGTGCTGTTTGGAGTAGGATATTTTAACTCAATAA	4920
Db	4861		
QY	4861	TTACTATTTATATGTAAGTACATGTGCTGTTTGGAGTAGGATATTTTAACTCAATAA	4920
Db	4861		
QY	4921	AGGTTATTTTCTTTTATTTGGGTTCAGGCAAGCTTCTTAAGGGGATGTGAAGGGATATCT	4980
Db	4921		
QY	4921	AGGTTATTTTCTTTTATTTGGGTTCAGGCAAGCTTCTTAAGGGGATGTGAAGGGATATCT	4980
Db	4921		
QY	4981	CTTTCTTTAGCTGAGGAGAGTGTAGTTCTTAAGTTAAATATATTAATCAAGGAATTTCCCT	5040
Db	4981		
QY	4981	CTTTCTTTAGCTGAGGAGAGTGTAGTTCTTAAGTTAAATATATTAATCAAGGAATTTCCCT	5040
Db	4981		
QY	5041	GTCTTTGCTATTTGAGATTTGTGACCAACAGCGGTTGGCTGAAAGGGAACTGAAGGG	5100
Db	5041		
QY	5041	GTCTTTGCTATTTGAGATTTGTGACCAACAGCGGTTGGCTGAAAGGGAACTGAAGGG	5100
Db	5041		
QY	5101	CGGGGAGGGGAGGAAATAGATGAAAAACAAAAACAAACTTCCCTTAAGCAGCTCT	5160
Db	5101		

Db	5101	CGGGGAGGAGGAAATAGATGAAAAACAAAAACAAATCTCCCTAAGCAGCTCT	5160
QY	5161	ACAAAACATTTTAGCCCCCAGAAATAGTCAAGAAATCTCAAAATCAAAACAGTATCCAGA	5220
Db	5161	ACAAAACATTTTAGCCCCCAGAAATAGTCAAGAAATCTCAAAATCAAAACAGTATCCAGA	5220
QY	5221	TACAAGGAAGTGTATGTAGCTCGAGCAGGGTGGACATCATCAGCTCAGTTTCAGTTACA	5280
Db	5221	TACAAGGAAGTGTATGTAGCTCGAGCAGGGTGGACATCATCAGCTCAGTTTCAGTTACA	5280
QY	5281	AAAGTCCAGGCTCTGAAATTAAGTCTGATGCCATTCATGCCAGCATCCAAATCAGACA	5340
Db	5281	AAAGTCCAGGCTCTGAAATTAAGTCTGATGCCATTCATGCCAGCATCCAAATCAGACA	5340
QY	5341	GAGATCAGAAAGTTCAGAGATGCCCTCCAGCTCCAAATTTGCCAAACAAAGTGTGGCTACTA	5400
Db	5341	GAGATCAGAAAGTTCAGAGATGCCCTCCAGCTCCAAATTTGCCAAACAAAGTGTGGCTACTA	5400
QY	5401	TACGTCAGAGGACTCTGAAGCCGTGAGAGGGGGAAGAACACAGTAGAGAGATGCCCA	5460
Db	5401	TACGTCAGAGGACTCTGAAGCCGTGAGAGGGGGAAGAACACAGTAGAGAGATGCCCA	5460
QY	5461	GCTGGTAAGATCGAGTGTATGAAGTTTATGATCAATTCATTCAGTCAATTCATTCAGTCA	5520
Db	5461	GCTGGTAAGATCGAGTGTATGAAGTTTATGATCAATTCATTCAGTCAATTCATTCAGTCA	5520
QY	5521	TCAAGAAACGCTCGGCTCTTTCGCAATATGATGAGGAGAGAGTGCCTAACTTCTA	5580
Db	5521	TCAAGAAACGCTCGGCTCTTTCGCAATATGATGAGGAGAGAGTGCCTAACTTCTA	5580
QY	5581	TGTCGTAGTACATTTGACCCCTATTTAGCTCCCGCTTTATATCTATATATATACAC	5640
Db	5581	TGTCGTAGTACATTTGACCCCTATTTAGCTCCCGCTTTATATCTATATATATACAC	5640
QY	5641	AGTATTTGTATATTTAT	5676
Db	5641	AGTATTTGTATATTTAT	5676
RESULT 5	BD093100	13940 bp DNA linear	PAT 27-AUG-2002
LOCUS	BD093100	Gene encoding promoter region of tumor suppressor gene p51.	
DEFINITION	BD093100	Gene encoding promoter region of tumor suppressor gene p51	
ACCESSION	BD093100	GI-22638688	
VERSION	BD093100.1	GI-22638688	
KEYWORDS	WO 0100818-A/12.	synthetic construct	
SOURCE	YOSHIOKAZU SUKENAGA, HIDEJI FUJII	other sequences; artificial sequences.	
ORGANISM	OS Artificial Sequence		
REFERENCE	1 (bases 1 to 13940)		
AUTHORS	Sakai, T., Kagaya, S., Sato, T., Sukenaga, Y. and Fujii, H.		
TITLE	Gene encoding promoter region of tumor suppressor gene p51		
JOURNAL	Patent: WO 0100818-A 12 04-JAN-2001;		
	TOSHIYUKI SAKAI, NIPPON KAYAKU KK, SHIGEHIDE KAGAYA, TAKAMICHI SATO,		
	YOSHIOKAZU SUKENAGA, HIDEJI FUJII		
COMMENT	OS Artificial Sequence		
	PN WO 0100818-A/12		
	PD 04-JAN-2001		
	PF 28-JUN-2000 WO 2000JP004261		
	PR 29-JUN-1999 JP 99 183195		
	PI TOSHIYUKI SAKAI, SHIGEHIDE KAGAYA, TAKAMICHI SATO, YOSHIOKAZU SUKENAGA,		
	PI HIDEJI FUJII		
	PC C12N15/10, C12N5/10, C12Q1/68, A61K48/00, A61K45/00, A61P43/00, PC A61P35/00		
	CC Plasmid gene containing p51promoter and neomycin resistance		
	CC gene		
	PH Key Location/Qualifiers.		
FEATURES	source		
	1. .13940		
	/organism="synthetic construct"		
	/mol type="genomic DNA"		
	/db_xref="taxon:32630"		

ORIGIN	Query Match	100.0%; Score 5673.4; DB 6; Length 13940;	
	Best Local Similarity	100.0%; Pred. No. 0;	
	Matches 5674; Conservative	0; Mismatches 1; Indels 0; Gaps 0;	
QY	2	AGCTGTTCAAGGATGTCTGGAAGAGAGCCACCCACCATTTGTTCTGGACACATGGGTGTG	61
Db	51	ATCTGTTTCAAGGATGTCTGGAAGAGAGCCACCCACCATTTGTTCTGGACACATGGGTGTG	110
QY	62	ACTTTGGAGGGTATCAGGTTTGTCTGTTAAAGAACTGCCAACTCTTCTCTGCCCAAT	121
Db	111	ACTTTGGAGGGTATCAGGTTTGTCTGTTAAAGAACTGCCAACTCTTCTCTGCCCAAT	170
QY	122	GGCCTCTGTTCCCTTGCATGCCCTCTTCTCTTGGGACACTCCCTTAAAGCATCTTCTTGA	181
Db	171	GGCCTCTGTTCCCTTGCATGCCCTCTTCTCTTGGGACACTCCCTTAAAGCATCTTCTTGA	230
QY	182	CATTAACTTAACATATAAATGTTTATTTGATGAATTTTCACTGACCTGAAGAGATGGAGG	241
Db	231	CATTAACTTAACATATAAATGTTTATTTGATGAATTTTCACTGACCTGAAGAGATGGAGG	290
QY	242	TCAAATCAGAAGACACATGCTAAGTTTGAATGCAATGCACTTGTCTTTTCAATTCAAATTA	301
Db	291	TCAAATCAGAAGACACATGCTAAGTTTGAATGCAATGCACTTGTCTTTTCAATTCAAATTA	350
QY	302	GTCATTTCGAATACCAATTCAGTTTACCTTAAAGTTCTAGGCCAGCTTTTACTCTTAATCGATG	361
Db	351	GTCATTTCGAATACCAATTCAGTTTACCTTAAAGTTCTAGGCCAGCTTTTACTCTTAATCGATG	410
QY	362	TCAGACTGTAGCAAAATATTTAGGTTCCAAAGTTGGAAGAGTTAGCAGGATCTCTCCATGAC	421
Db	411	TCAGACTGTAGCAAAATATTTAGGTTCCAAAGTTGGAAGAGTTAGCAGGATCTCTCCATGAC	470
QY	422	AGAACTTTGGCTTCCACTTTACTAAATAGAGATTTCTGTTGAGCTGAGCTAGCTATGAC	481
Db	471	AGAACTTTGGCTTCCACTTTACTAAATAGAGATTTCTGTTGAGCTGAGCTAGCTATGAC	530
QY	482	AGAAAGTGTACATACAAATTAATAAATCAACAACTCAGTCTCTTCAATTTGAGCAATAGTT	541
Db	531	AGAAAGTGTACATACAAATTAATAAATCAACAACTCAGTCTCTTCAATTTGAGCAATAGTT	590
QY	542	GGTGAATTTACTCCACACCTCTCTCTTGAAGTTCTTTTCTGCTCTCTCTCACTATAA	601
Db	591	GGTGAATTTACTCCACACCTCTCTCTTGAAGTTCTTTTCTGCTCTCTCTCACTATAA	650
QY	602	ATGCAGATGACCTGGAAGGCTAGCACTGAGTTTCACTGAGTTTACCTGACACAAAGGAAT	661
Db	651	ATGCAGATGACCTGGAAGGCTAGCACTGAGTTTCACTGAGTTTACCTGACACAAAGGAAT	710
QY	662	CAGTTTCTCTGATCTCATAGTCAAGGCTGCCAGAGCTCTACGGAAACATGCAAGATCATC	721
Db	711	CAGTTTCTCTGATCTCATAGTCAAGGCTGCCAGAGCTCTACGGAAACATGCAAGATCATC	770
QY	722	TGCTTTAAGCCTCTTGTGGGCAATCTGTTTTCACCTGAGCTGACCTTATGCTCTT	781
Db	771	TGCTTTAAGCCTCTTGTGGGCAATCTGTTTTCACCTGAGCTGACCTTATGCTCTT	830
QY	782	TCCTGGTTTAAAGAACCTTTTATTTCTGTAAGTCTCTGCTGAGCTATGCTAGGCTG	841
Db	831	TCCTGGTTTAAAGAACCTTTTATTTCTGTAAGTCTCTGCTGAGCTATGCTAGGCTG	890
QY	842	CATCAGTCCACATGATCAGGCTCTCTGCGCAAAACATGGCATCTTCTTTTGGGAATTT	901
Db	891	CATCAGTCCACATGATCAGGCTCTCTGCGCAAAACATGGCATCTTCTTTTGGGAATTT	950
QY	902	GAATCTTAAAGTGAATAGCTGAAGTTTCAAAAAAGCTGTTGAATCTGACCTACGCTTACA	961
Db	951	GAATCTTAAAGTGAATAGCTGAAGTTTCAAAAAAGCTGTTGAATCTGACCTACGCTTACA	1010
QY	962	GTGGCTTGAAGAGTGAATAGCTGCAATCTTCTTATTTCTTAAAGTCCCTGAAATTTATTTATC	1021
Db	1011	GTGGCTTGAAGAGTGAATAGCTGCAATCTTCTTATTTCTTAAAGTCCCTGAAATTTATTTATC	1070

Qy	1022	CTGGTTACAGCCCTTCTGAGATGTGGTGTCTTTTTCCTTCTTATAGTCTGTG	1081
Db	1071	CTGGTTACAGCCCTTCTGAGATGTGGTGTCTTTTTCCTTCTTATAGTCTGTG	1130
Qy	1082	AAATTTTCATATTTCTTTTTCATACATTTTCAATTTTCTTGTGTTGTGTTTGTG	1141
Db	1131	AAATTTTCATATTTCTTTTTCATACATTTTCAATTTTCTTGTGTTGTGTTTGTG	1190
Qy	1142	GCATTAGGTAGGAGATCAGTTCTGTTCTGTTTATATACCAAGGAATCCTGATTGATCAT	1201
Db	1191	GCATTAGGTAGGAGATCAGTTCTGTTCTGTTTATATACCAAGGAATCCTGATTGATCAT	1250
Qy	1202	CTTTCCCTTTTAAAAATAAAGTATCTAAGGCTCAAGAGAGTAGGCTACCTGCTGAGGT	1261
Db	1251	CTTTCCCTTTTAAAAATAAAGTATCTAAGGCTCAAGAGAGTAGGCTACCTGCTGAGGT	1310
Qy	1262	CTGGGAGTAAGTTAGTACCAAGAGCTCGTACTAAACCCAGGTTAGCCAACTGCTTTTACACA	1321
Db	1311	CTGGGAGTAAGTTAGTACCAAGAGCTCGTACTAAACCCAGGTTAGCCAACTGCTTTTACACA	1370
Qy	1322	ACATTTGCTCTCTCTTTCAGAGTTATAGCAGTCTTGGGAAGAAAGAGCTACTATTTTGCC	1381
Db	1371	ACATTTGCTCTCTCTTTCAGAGTTATAGCAGTCTTGGGAAGAAAGAGCTACTATTTTGCC	1430
Qy	1382	AAAGACCTCAGGAGGACCAAGAACAGTCTCTGGGATATGTGATTTGAACCTCTTAAAAA	1441
Db	1431	AAAGACCTCAGGAGGACCAAGAACAGTCTCTGGGATATGTGATTTGAACCTCTTAAAAA	1490
Qy	1442	GTTTGTTGGACTTTCTGCGCCATAATTTGTGTATCTAAGACCAAGATTTCAATCTTAAATAGCTA	1501
Db	1491	GTTTGTTGGACTTTCTGCGCCATAATTTGTGTATCTAAGACCAAGATTTCAATCTTAAATAGCTA	1550
Qy	1502	AACAAACAAACAGAGATCCACAGGTTACAGAGCTATAATAAGAGTGAATTTACTGTATACA	1561
Db	1551	AACAAACAAACAGAGATCCACAGGTTACAGAGCTATAATAAGAGTGAATTTACTGTATACA	1610
Qy	1562	GTTGACAAATGAATATATCTCAGAACCAATGGCATCAATGAGCAAAAAAATCCAGACA	1621
Db	1611	GTTGACAAATGAATATATCTCAGAACCAATGGCATCAATGAGCAAAAAAATCCAGACA	1670
Qy	1622	CAGAAGAAATACGTAACATATGCTGCAATTTATGTGATTTCTAGCATTTGATTTGCCAAC	1681
Db	1671	CAGAAGAAATACGTAACATATGCTGCAATTTATGTGATTTCTAGCATTTGATTTGCCAAC	1730
Qy	1682	ATAGTAGCCAGTAGCCCTACATGCGCTATTCAAATTTAAGTTGATTAATAATTAAGTAAGAA	1741
Db	1731	ATAGTAGCCAGTAGCCCTACATGCGCTATTCAAATTTAAGTTGATTAATAATTAAGTAAGAA	1790
Qy	1742	TAAAAATTTTAGCTCTTCAGTAGGTTAGCCACATGTAACCTAGTGGCTACCACATCAGACG	1801
Db	1791	TAAAAATTTTAGCTCTTCAGTAGGTTAGCCACATGTAACCTAGTGGCTACCACATCAGACG	1850
Qy	1802	GTGCAAAATATAGAAATTTCTTTTATACAGAAAGTTCTATTGGAAAAAATGTTCTAGA	1861
Db	1851	GTGCAAAATATAGAAATTTCTTTTATACAGAAAGTTCTATTGGAAAAAATGTTCTAGA	1910
Qy	1862	AAATATACATATCTATAAAAAAGCAAGTCAAGTCAAGTGTGTTCTAAGGCCAGGGGT	1921
Db	1911	AAATATACATATCTATAAAAAAGCAAGTCAAGTCAAGTGTGTTCTAAGGCCAGGGGT	1970
Qy	1922	GAGGGAGATCGATTGCAAAAGTGTATGAGGAAGTTTTCGGGTAAATAGGGTTGTTGGAA	1981
Db	1971	GAGGGAGATCGATTGCAAAAGTGTATGAGGAAGTTTTCGGGTAAATAGGGTTGTTGGAA	2030
Qy	1982	TCATTGATTGGAGTGAAGGCTACTCGGTGTCTAATGTGTCACTCTCAGACTGAAACACTT	2041
Db	2031	TCATTGATTGGAGTGAAGGCTACTCGGTGTCTAATGTGTCACTCTCAGACTGAAACACTT	2090
Qy	2042	GGAATTTGGCGAAATTTCAATTGTATGTAATTTATACCTCATTAAGTAACTCTAAGAGGTCAA	2101
Db	2091	GGAATTTGGCGAAATTTCAATTGTATGTAATTTATACCTCATTAAGTAACTCTAAGAGGTCAA	2150

Qy	2102	GTCTTTTGTGGAAATTAATTTTAAATCAGTTGCAATACTTATTATGAGATGATTTTGGCAA	2161
Db	2151	GTCTTTTGTGGAAATTAATTTTAAATCAGTTGCAATACTTATTATGAGATGATTTTGGCAA	2210
Qy	2162	ATACATAAACATGTTTATTCATTCATTTAGGTGCAATATTTTGTCTAGCTCTGAAAAACACA	2221
Db	2211	ATACATAAACATGTTTATTCATTCATTTAGGTGCAATATTTTGTCTAGCTCTGAAAAACACA	2270
Qy	2222	GAGATCAATTAGAATAGCAAGCTGCCCTCAAGCTCTTCAATCCAGTACAGTACAGAGATGA	2281
Db	2271	GAGATCAATTAGAATAGCAAGCTGCCCTCAAGCTCTTCAATCCAGTACAGTACAGAGATGA	2330
Qy	2282	GTCTATTCAAAAATAGCTAGACTCCAGAGAAAGAAAGTTATAGGTGACCTTACACAAAAAAG	2341
Db	2331	GTCTATTCAAAAATAGCTAGACTCCAGAGAAAGAAAGTTATAGGTGACCTTACACAAAAAAG	2390
Qy	2342	TGCAGATATAATTTATGTAGGACAGTAGAAGTGGGAAAGTTTCTTTTATGTGAAAAAAG	2401
Db	2391	TGCAGATATAATTTATGTAGGACAGTAGAAGTGGGAAAGTTTCTTTTATGTGAAAAAAG	2450
Qy	2402	AGGGAGAAATTTTGGTCTTTTGAAGGATGACAGATGTGAATATGCCAGATGCGATTTT	2461
Db	2451	AGGGAGAAATTTTGGTCTTTTGAAGGATGACAGATGTGAATATGCCAGATGCGATTTT	2510
Qy	2462	AAAAACATTCCTGCTGAGGCGCAGAAATATGATCCAAGGCAACAGAGCAACCCAGAAAAATAT	2521
Db	2511	AAAAACATTCCTGCTGAGGCGCAGAAATATGATCCAAGGCAACAGAGCAACCCAGAAAAATAT	2570
Qy	2522	GCAACCTAGAGGAAAGTGCATGAAGGGGAGCTAGTTGTAAAAATAATTTTCATGAATGTAAAG	2581
Db	2571	GCAACCTAGAGGAAAGTGCATGAAGGGGAGCTAGTTGTAAAAATAATTTTCATGAATGTAAAG	2630
Qy	2582	TGAGAGAATTTTGTATCATAGACACCTGAGTTTGGCAGAGTGCATGTTCTTGGCTCCTAG	2641
Db	2631	TGAGAGAATTTTGTATCATAGACACCTGAGTTTGGCAGAGTGCATGTTCTTGGCTCCTAG	2690
Qy	2642	GAGTCAAGAGAAACAAAGTGTCCCTTCTCCTACAGTTATGCTCAGTGGTCCAAAGTCCAAA	2701
Db	2691	GAGTCAAGAGAAACAAAGTGTCCCTTCTCCTACAGTTATGCTCAGTGGTCCAAAGTCCAAA	2750
Qy	2702	ACACCTTTCTCTTCTTAAAGTACTTTCTCTCCCTCCATACAAATCTAAAGTCTTCACAA	2761
Db	2751	ACACCTTTCTCTTCTTAAAGTACTTTCTCTCCCTCCATACAAATCTAAAGTCTTCACAA	2810
Qy	2762	ACATCAATTTAAACAGCAGGTCATGCTCAGAAAGCAATGCTTTTCTTAGACTTCTATG	2821
Db	2811	ACATCAATTTAAACAGCAGGTCATGCTCAGAAAGCAATGCTTTTCTTAGACTTCTATG	2870
Qy	2822	TAGCTTATATATTAACAAATTTCTGCCCTAAAAAGACTCTAAAGTCTTGGAAAAAGTTTCCACC	2881
Db	2871	TAGCTTATATATTAACAAATTTCTGCCCTAAAAAGACTCTAAAGTCTTGGAAAAAGTTTCCACC	2930
Qy	2882	TTGCACATCAAGAGATATAATTCATGCAATTTGTATAGTAACCTTTAGTCCCTTAAGAAATA	2941
Db	2931	TTGCACATCAAGAGATATAATTCATGCAATTTGTATAGTAACCTTTAGTCCCTTAAGAAATA	2990
Qy	2942	AGGATCAATTAATAATTAAGAGTAATTTATGTAATTAATAATGATTCGCCACTTATTT	3001
Db	2991	AGGATCAATTAATAATTAAGAGTAATTTATGTAATTAATAATGATTCGCCACTTATTT	3050
Qy	3002	TTCACTTTGATCGGTATGTTGTCATGCTGCTGTTCTGTTGAAATTTCTAGAGAGTTTGC	3061
Db	3051	TTCACTTTGATCGGTATGTTGTCATGCTGCTGTTCTGTTGAAATTTCTAGAGAGTTTGC	3110
Qy	3062	CTCTTTTCTCGGGTCAACTCGCCATTTATTTCCATTAATGCAATAGGAGCCAACTCTTT	3121
Db	3111	CTCTTTTCTCGGGTCAACTCGCCATTTATTTCCATTAATGCAATAGGAGCCAACTCTTT	3170
Qy	3122	TTCAATAATTAATTTTAAATTTTGTGCAATTTAATTTCTGTTCTCTTAGCTTAGTAA	3181
Db	3171	TTCAATAATTAATTTTAAATTTTGTGCAATTTAATTTCTGTTCTCTTAGCTTAGTAA	3230
Qy	3182	CTTTAGGATTTTAAATAACAACTATTGAAATCATGACATACGTTTTAAATGATATTATTT	3241

Db 3231 |||||CTTTAGGATTTTAAATAAACAACCTATTGAATCATGACATACGTTTAAATGATATATTT 3290
Qy 3242 AATAAGTTAGGCTATAAAGCTTTTAAATTTTAAATAATAGTAGTGCTGGCTC 3301
Db 3291 AATAAGTTAGGCTATAAAGCTTTTAAATTTTAAATAATAGTAGTGCTGGCTC 3350
Qy 3302 ATCCCTGTATATCCCAACACCTTTTGGGAAGCCGGGTCCGGAGGATAGCTTGAAGTCCAGCAGT 3361
Db 3351 ATGCTGTATATCCCAACACCTTTTGGGAAGCCGGGTCCGGAGGATAGCTTGAAGTCCAGCAGT 3410
Qy 3362 TTGAGACAGTACAGGCAACACAGCAAGACCCCATATCTTAAATAAACAACAAACAAA 3421
Db 3411 TTGAGACAGTACAGGCAACACAGCAAGACCCCATATCTTAAATAAACAACAAACAAA 3470
Qy 3422 ATTTACCTGGGTATGTTGCTCTCACCTGTAGTCCAGCTACACAGGAAGCTGAGGAGAA 3481
Db 3471 ATTTACCTGGGTATGTTGCTCTCACCTGTAGTCCAGCTACACAGGAAGCTGAGGAGAA 3530
Qy 3482 GGATCACTTGAGCCAGGAGGTTGAGGCTGAGCTGATCCATGAACGGCTGCTACACTCA 3541
Db 3531 GGATCACTTGAGCCAGGAGGTTGAGGCTGAGCTGATCCATGAACGGCTGCTACACTCA 3590
Qy 3542 GTCTGGGTGACGTGCAAGAGCTGCTCAAAATAATAATAATAATAATAATAATAATAATA 3601
Db 3591 GTCTGGGTGACGTGCAAGAGCTGCTCAAAATAATAATAATAATAATAATAATAATAATA 3650
Qy 3602 AAAAAAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 3661
Db 3651 AAAAAAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 3710
Qy 3662 TATTTGGGAGCTATATCCCTGGAGGTTAATTTAAATAATAATAATAATAATAATAATAATA 3721
Db 3711 TATTTGGGAGCTATATCCCTGGAGGTTAATTTAAATAATAATAATAATAATAATAATAATA 3770
Qy 3722 TTCTTAGAGACGTGCAATTTGTAATATCAGAGCTAGAGGAACAATAAGGCTCGCCACTC 3781
Db 3771 TTCTTAGAGACGTGCAATTTGTAATATCAGAGCTAGAGGAACAATAAGGCTCGCCACTC 3830
Qy 3782 CAAAGTGTGTCNAGACACAGCAGCATCAAGTAACCTGGGAAGCTGTTAGAATGCAGA 3841
Db 3831 CAAAGTGTGTCNAGACACAGCAGCATCAAGTAACCTGGGAAGCTGTTAGAATGCAGA 3890
Qy 3842 GTCTTAGGCTCACCACAGACCTACTCAACAGAGCTGCAATTAACAAGATTTCTAGGTG 3901
Db 3891 GTCTTAGGCTCACCACAGACCTACTCAACAGAGCTGCAATTAACAAGATTTCTAGGTG 3950
Qy 3902 CCTCACGGGCACATTAATACTTGAGAAAGCTCTGCACTAGAAATCTTCACTCCACCTTTCA 3961
Db 3951 CCTCACGGGCACATTAATACTTGAGAAAGCTCTGCACTAGAAATCTTCACTCCACCTTTCA 4010
Qy 3962 TTATAAATGAATCACTTGGGCTGTGTGTCACAGGAATGATATTTTAAATTTTCAAGAC 4021
Db 4011 TTATAAATGAATCACTTGGGCTGTGTGTCACAGGAATGATATTTTAAATTTTCAAGAC 4070
Qy 4022 CTTCATTTTAGGCTCATCTATTTTGTCTAATAGCAGGAAGAACCCAACTCTTTAACTG 4081
Db 4071 CTTCATTTTAGGCTCATCTATTTTGTCTAATAGCAGGAAGAACCCAACTCTTTAACTG 4130
Qy 4082 CAATTAACAATACTATAAATTAATTAAGCAATCTTCCCTTTAAGTTTAACTTTTACATTTGT 4141
Db 4131 CAATTAACAATACTATAAATTAATTAAGCAATCTTCCCTTTAAGTTTAACTTTTACATTTGT 4190
Qy 4142 GGAGCAAGCTGTTTGAATTTGGCTGGGCTCAGGCGGCTGTTTGTGAATTTCAAAATTC 4201
Db 4191 GGAGCAAGCTGTTTGAATTTGGCTGGGCTCAGGCGGCTGTTTGTGAATTTCAAAATTC 4250
Qy 4202 ACAGATCTTAGCCCTCTCGGCTTAAGTAAAGGAGAGAGTGTCAAGTTTAAATAGCTT 4261
Db 4251 ACAGATCTTAGCCCTCTCGGCTTAAGTAAAGGAGAGAGTGTCAAGTTTAAATAGCTT 4310
Qy 4262 CTCCCTTCCATCTCGGCTGAAGCAACAATAAATAATTTTATGAACAACATTTTGAAGTT 4321

Db 4311 CTCCCTTCCATCTCGCTGAAGCAACAATAAATAATTTTATGAACAACATTTTGAAGTT 4370
Qy 4322 AGATTTACTTTACAGGAAATGTCAAAATTTCTCTGAAGGGCTTTAGATTTGTCTCAAACT 4381
Db 4371 AGATTTACTTTACAGGAAATGTCAAAATTTCTCTGAAGGGCTTTAGATTTGTCTCAAACT 4430
Qy 4382 TTGACATCTACATGATGTCACCTATTTTACAGGGTGTCTGTGACTAGGGGGTGAAGGAA 4441
Db 4431 TTGACATCTACATGATGTCACCTATTTTACAGGGTGTCTGTGACTAGGGGGTGAAGGAA 4490
Qy 4442 GATGTGAATCTCAACATGTTAGTGACCGTTAGATACACAGAGTGGTTTTTTTCCCTCTGT 4501
Db 4491 GATGTGAATCTCAACATGTTAGTGACCGTTAGATACACAGAGTGGTTTTTTTCCCTCTGT 4550
Qy 4502 TGGAGTCTATCTCTAACTGAGCTTCTGAATCATATTTCAATTTCCAAATCCACAAA 4561
Db 4551 TGGAGTCTATCTCTAACTGAGCTTCTGAATCATATTTCAATTTCCAAATCCACAAA 4610
Qy 4562 CCAGGATAAGTTTACAGCCCATNTCAGAAAGGAATAATAATTTTGTGTGTAGACTTT 4621
Db 4611 CCAGGATAAGTTTACAGCCCATNTCAGAAAGGAATAATAATTTTGTGTGTAGACTTT 4670
Qy 4622 CCTGATATTACACTGATTTGGGAATATATGAACAATTTTATGGTTTCTTCCGAAGTAGG 4681
Db 4671 CCTGATATTACACTGATTTGGGAATATATGAACAATTTTATGGTTTCTTCCGAAGTAGG 4730
Qy 4682 TCAAGTCAAAGCAAAACCAAAACAGCAAAACCTGTAAGACATAAAGAAATAGAGTGGAGC 4741
Db 4731 TCAAGTCAAAGCAAAACCAAAACAGCAAAACCTGTAAGACATAAAGAAATAGAGTGGAGC 4790
Qy 4742 CGACTGAGAGATTAATAATAAATACTAGAAATATTTTATTAACAGGCAATTTGAAATAATTG 4801
Db 4791 CGACTGAGAGATTAATAATAAATACTAGAAATATTTTATTAACAGGCAATTTGAAATAATTG 4850
Qy 4802 TGCACTTCAGATATTCTACAATAATATATTTTCCAAATTTTAAATCTTTTAAAGAAAT 4861
Db 4851 TGCACTTCAGATATTCTACAATAATATATTTTCCAAATTTTAAATCTTTTAAAGAAAT 4910
Qy 4862 TACTATATTATATGTAAGTACATGTGCATGTGTTTGAGGTAGGATATTTAACTCAATAA 4921
Db 4911 TACTATATTATATGTAAGTACATGTGCATGTGTTTGAGGTAGGATATTTAACTCAATAA 4970
Qy 4922 GGTATTTTCTTTTATTTCCGGGTGAGGCAAGCTTCTAAGGGGATGTGAAGGAGATATCTC 4981
Db 4971 GGTATTTTCTTTTATTTCCGGGTGAGGCAAGCTTCTAAGGGGATGTGAAGGAGATATCTC 5030
Qy 4982 TTTCTCTTAGCTGAGAGGAGTGTGATTTCTAAGTTTAAATAATAATACTAAGGAATTTCCCTG 5041
Db 5031 TTTCTCTTAGCTGAGAGGAGTGTGATTTCTAAGTTTAAATAATAATACTAAGGAATTTCCCTG 5090
Qy 5042 TCTTTGCTATTTGAGATTTGACCAACAGGCGGTTGGCTGAAGGGAACCTGAAGGGC 5101
Db 5091 TCTTTGCTATTTGAGATTTGACCAACAGGCGGTTGGCTGAAGGGAACCTGAAGGGC 5150
Qy 5102 GGGGAGGAGGGAATAGATGAATAAACAACAAACAAACAAACAAACAAACAAACAAACAA 5161
Db 5151 GGGGAGGAGGGAATAGATGAATAAACAACAAACAAACAAACAAACAAACAAACAAACAA 5210
Qy 5162 CAAACAAATTTTAGCCCCCAGAAATAGTCAAGAAATCTCTCAATCAAAACAGATCCAGAT 5221
Db 5211 CAAACAAATTTTAGCCCCCAGAAATAGTCAAGAAATCTCTCAATCAAAACAGATCCAGAT 5270
Qy 5222 ACAGGAGGTTTATGTAGCTGAGAGGAGTGTGACACTCATCAGCTCAGTTCAAGTTTCAAA 5281
Db 5271 ACAGGAGGTTTATGTAGCTGAGAGGAGTGTGACACTCATCAGCTCAGTTCAAGTTTCAAA 5330
Qy 5282 AAGTCCAGGCTGTGAAATTAATACTCTGATGCAATTTATGTCAGCATCCAAATCAGCAG 5341
Db 5331 AAGTCCAGGCTGTGAAATTAATACTCTGATGCAATTTATGTCAGCATCCAAATCAGCAG 5390
Qy 5342 AGATCAGAGTTTCAAGATGCTCCAGCTCCTCAATTTGCGCAACAAAGTGTGGCTACTAT 5401
Db 5391 AGATCAGAGTTTCAAGATGCTCCAGCTCCTCAATTTGCGCAACAAAGTGTGGCTACTAT 5450

Qy	5402	ACGTCAAGACTCTGAAGCCGTCGAGAGAGGGGGAAGAACAACAGTAGAGAGATGCCCAG	5461
Dd	5451	ACGTCAAGGACTCTGAAGCCGTCGAGAGAGGGGGAAGAACAACAGTAGAGAGATGCCCAG	5510
Qy	5462	CTCGTAGAATCGAGTGTCTTTATGAAGTTTTAGTCAATTGATGAATCTCATTTGCTTAATAAT	5521
Dd	5511	CTGGAAGAATCGAGTGTCTTTATGAAGTTTTAGTCAATTGATGAATCTCATTTGCTTAATAAT	5570
Qy	5522	CAAGAAACGCTCCGCCTCTTTTGCAAATATATGATGAAGGAGAGAGTGCCTAAAACCTTCTAT	5581
Dd	5571	CAAGAAACGCTCCGCCTCTTTTGCAAATATATGATGAAGGAGAGAGTGCCTAAAACCTTCTAT	5630
Qy	5582	GTCGTATGATTTGACCCTATTGCTTTTATGCTCCCGCTTTATATCTATATATATACACA	5641
Dd	5631	GTCGTATGATTTGACCCTATTGCTTTTATGCTCCCGCTTTATATCTATATATATACACA	5690
Qy	5642	GGTATTGTCGTATATTTTATATATATTTCTTCCTCGGT	5676
Dd	5691	GGTATTGTCGTATATTTTATATATATTTCTTCCTCGGT	5725
RESULT	6		
LOCUS	AC078809/c		
DEFINITION	Homo sapiens 3 BAC RP1-600G3 (Roswell Park Cancer Institute Human BAC library) complete sequence.	linear	PRI 25-JAN-2003
ACCESSION	AC078809		
VERSION	AC078809.9	GI:190333390	
KEYWORDS	HTG.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 162546)		
AUTHORS	Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C., Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Ayele,M., Banks,T., Barbara,J., Benton,J., Binage,K., Blankenburg,K., Bonnin,D., Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C., Burck,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Deigado,O., Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., He,X., Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M., Hollway,C., Hollins,B., Homai,F., Howard,S., Huber,J.S., Huliyk,S., Hume,J., Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Joliviet,S., Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L., Li,J., Li,Z., Lichtarge,O., Lied,C., Liu,J., Liu.W., Loulseghe,H., Lozando,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E., Massey,E., Mawhinney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Moore,S., Morgan,M., Moorish,T., Morris,S., Moser,M., Neal,D., Nelson,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokwenkwo,S., Ogih,M., Okunishi,G., Ornguyen,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L., Peters,L., Pickens,R., Primm,E., Pu,L.I., Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojubenok,I., Rolfe,M., Ruiz,S., Savory,G., Scherer,S., Scott,G., Shen,H., Shooshbari,N., Sisson,I., Sodergren,E., Sonake,T., Sparks,A., Stanley,H., Stone,H., Sutton,A., Svatek,A., Taborski,P., Tamerisa,A., Tamerisa,K., Tang,H., Tansey,J., Taylor,C., Taylor,T., Teifrod,B., Thomas,N., Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q., Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S., Williams,G., Williams,G.,		

Williamson, A., Wleczyk, R., Wooden, S., Worley, K., Wu, C., Wu, Y., Wu, Y. F., Zhou, J., Zorrilla, S., Naylor, S. L., Weinstock, G. and Gibbs, R.

Direct Submission
Unpublished
2 (bases 1 to 162646)
Worley, K.C.

Direct Submission
Submitted (04-AUG-2000) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 162646)
Worley/K.C.
Direct Submission
Submitted (01-MAR-2002) Human Genome Sequencing Center, Department

of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
4 (bases 1 to 162646)
Worley, K.C.

Direct Submission
Submitted (29-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

5 (bases 1 to 162646)
 Worley, K.C.
 Direct Submission
 Submitted (30-AUG-2002) Human Genome Sequencing Center, Department
 of Molecular Biology and Genetics, Baylor College of Medicine, One

Department of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
6 (bases 1 to 162646)
Worley, K.C.
Direct Submission

Submitted (25-JAN-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On Mar 1 2003 this sequence version replaced at:17977513

CLONE LENGTH: This sequence does not necessarily represent the
 INFORMATION: <http://www.hgsc.bcm.tmc.edu/> or email
gc-help@bcm.tmc.edu

entire insert of this clone. Overlapping regions of clones are only sequenced and submitted once, so the sequence for the remainder of the insert may be found in the record for the adjacent clones. Overlapping clones are noted at the beginning and end of the

Features listing.

ANNOTATION OF FEATURES:
STSS are identified using ePCR (Genome Res. 7:541-550) searches

Repeats are identified using RepeatMasker (A. Smit and P. Green unpublished.) for Human and Mouse sequences.

Genes and region of sequence similarity are identified by BLAST (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the EST and cDNA sequences. Genes demonstrate at least two exons flanked by consensus splice sites that maintained sequence

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum of 100-fold with minimum overlap of 100 bp.

reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality for a region does not meet this standard, it will be indicated in the

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base annotation as low coverage.

reports of lowest quality, individual bases and measures of base quality are listed below. Description of the metrics can be found at URL: <http://gc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html>.

QUALSTAT-REPORT.		
FEATURES	Location/Qualifiers	
source	1. .162646	
	/organism="Homo sapiens"	
	/mol_type="genomic DNA"	
	/db_xref="taxon:9606"	
	/chromosome="3"	
	/clone="RP11-600G3"	
	1. .2004	
misc_feature	/notes="overlaps bases 79511. .81510 of clone AC117486"	
	/function="clone overlap"	
repeat_region	complement(248. .339)	
repeat_region	/rpt_family="MIR"	
repeat_region	491. .556	
repeat_region	/rpt_family="TA)n"	
repeat_region	complement(844. .981)	
repeat_region	/rpt_family="THE1B"	
repeat_region	1586. .1727	
repeat_region	/rpt_family="MIR"	
repeat_region	1683. .1736	
repeat_region	/rpt_family="L2"	
repeat_region	2734. .2922	
repeat_region	/rpt_family="MIR"	
repeat_region	2977. .3083	
repeat_region	/rpt_family="L2"	
repeat_region	complement(3383. .3562)	
repeat_region	/rpt_family="MIR"	
repeat_region	3742. .7012	
repeat_region	/rpt_family="L1MC3"	
repeat_region	7013. .7275	
repeat_region	/rpt_family="AluSx"	
repeat_region	7276. .7892	
repeat_region	/rpt_family="L1MC3"	
repeat_region	complement(8146. .8365)	
repeat_region	/rpt_family="MIR"	
repeat_region	9447. .9479	
repeat_region	/rpt_family="AT-rich"	
repeat_region	9648. .9780	
repeat_region	/rpt_family="GA-rich"	
repeat_region	10440. .10660	
repeat_region	/rpt_family="MER20"	
repeat_region	complement(10703. .11014)	
repeat_region	/rpt_family="AluSx"	
STS	11431. .11683	
	/standard_name="79189"	
repeat_region	complement(11911. .12013)	
repeat_region	/rpt_family="MIR"	
repeat_region	12854. .12964	
repeat_region	/rpt_family="AluSx"	
repeat_region	13022. .13142	
repeat_region	/rpt_family="MIR"	
repeat_region	15041. .15337	
repeat_region	/rpt_family="AluSg"	
repeat_region	15456. .15480	
repeat_region	/rpt_family="TTTA)n"	
repeat_region	complement(16012. .16151)	
repeat_region	/rpt_family="MIR"	
repeat_region	17138. .17449	
repeat_region	/rpt_family="AluSp"	
repeat_region	17765. .17895	
repeat_region	/rpt_family="L2"	
repeat_region	18634. .18860	
repeat_region	/rpt_family="MIR"	
repeat_region	18907. .19141	
repeat_region	/rpt_family="MIR"	
repeat_region	19474. .19515	
repeat_region	/rpt_family="(TA)n"	
repeat_region	complement(19940. .20118)	
repeat_region	/rpt_family="AluSx"	
repeat_region	20726. .21393	
repeat_region	/rpt_family="L1M3e"	
repeat_region	complement(21402. .21704)	
repeat_region	/rpt_family="L1"	
repeat_region	21710. .23051	
repeat_region	/rpt_family="L1M2"	
repeat_region	23048. .24606	
repeat_region	/rpt_family="L1M1"	
repeat_region	24611. .24765	
Query Match	99.3%; Score 5635.6; DB 9; Length 162646;	
Best Local Similarity	99.8%; Pred. No. 0;	
Matches 5666; Conservative	0; Mismatches 9; Indels 5; Gaps 2;	
QY	1 CAGCTCTTCAGGATGCTCTGGAAAGAACCCACCCACACATTGCTTCTGGACACTGGGTGT	60
Db		
QY	76260 CAGCTGTTTCAGGATGCTCTGGAAAGAACCCACCCACACATTGCTTCTGGACACTGGGTGT	76201
Db		
QY	61 GACTTTGGAGGGTATCAGGTTTGTCTGTTAAAGAAATGCAACCTCTTCTCCGCCCAAT	120
Db		
QY	76200 GACTTTGGAGGGTATCAGGTTTGTCTGTTAAAGAAATGCAACCTCTTCTCCGCCCAAT	76141
Db		
QY	121 TGGCCTCTGTTCCCTTGCATGCCCTCTTTCCTTGGGACACTCCCTTAAGGCATCTTCTTG	180
Db		
QY	76140 TGGCCTCTGTTCCCTTGCATGCCCTCTTTCCTTCGGACACTCCCTTAAGGCATCTTCTTG	76081
Db		
QY	181 ACATTAACTTAACATAAATGTTTATTTGATGAATTTTCAGTGACCTGAAGAGAGATGGAG	240
Db		
QY	76080 ACATTAACTTAACATAAATGTTTATTTGATGAATTTTCAGTGACCTGAAGAGAGATGGAG	76021
Db		
QY	241 GTCAAAATCAGAAGACACATGCTAAGGTTGCAATGCACCTTCTTTTCAATTGAATTA	300
Db		
QY	76020 GTCAAAATCAGAAGACACATGCTAAGGTTGCAATGCACCTTCTTTTCAATTGAATTA	75961
Db		
QY	301 AGTCATTTCGAATACCAATTCAGTTTACTTAAAGTTCTAGGCCCAGCTTCTCTTAATCGAT	360
Db		
QY	75960 AGTCATTTCGAATACCAATTCAGTTTACTTAAAGTTCTAGGCCCAGCTTCTCTTAATCGAT	75901
Db		
QY	361 GTCAGACTGTAGCAATATTTAGTCCAAAGTTTCGAAGAGTTAGCAGGATCCTCTCCATGA	420
Db		
QY	75900 GTCAGACTGTAGCAATATTTAGTCCAAAGTTTCGAAGAGTTAGCAGGATCCTCTCCATGA	75841
Db		
QY	421 CAGAACTTTGGCTTCCACTTTTACTTAAATAGAGATTGTGTGGTTGAGCTGCAGCTATGTA	480
Db		
QY	75840 CAGAACTTTGGCTTCCACTTTTACTTAAATAGAGATTGTGTGGTTGAGCTGCAGCTATGTA	75781
Db		
QY	481 CAGAAAGTGTCAATAATTTAAATACCAAACTCAGTCTCTTCAATTTGAGCAATAGT	540
Db		
QY	75780 CAGAAAGTGTCAATAATTTAAATACCAAACTCAGTCTCTTCAATTTGAGCAATAGT	75721
Db		
QY	541 TGGTGAATTTACTCCACCACCTCTCTCTTGAAGTTCTTCTCTGCTCTCCCTCATATA	600
Db		
QY	75720 TGGTGAATTTACTCCACCACCTCTCTCTTGAAGTTCTTCTCTGCTCTCCCTCATATA	75661
Db		
QY	601 AATGCAGGATGACCTGGAAAGGCTTAGGACCTGAGGTTTCAGTTACCTTGACAAAGGAAT	660
Db		
QY	75660 AATGCAGGATGACCTGGAAAGGCTTAGGACCTGAGGTTTCAGTTACCTTGACAAAGGAAT	75601
Db		
QY	661 TCAGTTTCTCTGATCTCATAGTCACAGGCTGCCAGAGCTCTACGGAAACATGCAAGATCAT	720
Db		
QY	75600 TCAGTTTCTCTGATCTCATAGTCACAGGCTGCCAGAGCTCTACGGAAACATGCAAGATCAT	75541
Db		
QY	721 CTGCTTTAAGCTCTTGTGTGGCATCTGTGTTTCCACTGCCCTGCTACTATTGCTCT	780
Db		
QY	75540 CTGCTTTAAGCTCTTGTGTGGCATCTGTGTTTCCACTGCCCTGCTACTATTGCTCT	75481
Db		
QY	781 TTCTCTGGTTTAACAGAACCTTTATTTTCTTCTGAAAACTCTGCTCAGTCAATGGTAGGG	840
Db		
QY	75480 TTCTCTGGTTTAACAGAACCTTTATTTTCTTCTGAAAACTCTGCTCAGTCAATGGTAGGG	75421
Db		
QY	841 CCATCAGTCCACATGATCAGGCTCTCTCTGGCCAAACATGGCATCTTTCTTTGGGAATT	900
Db		
QY	75420 CCATCAGTCCACATGATCAGGCTCTCTCTGGCCAAACATGGCATCTTTCTTTGGGAATT	75361
Db		
QY	901 TGAATCTTAAGCTGAATAGCTGAAGTTCAAAAAAGCTGTTGAATCTGACTTACGCCTAC	960
Db		
QY	75360 TGAATCTTAAGCTGAATAGCTGAAGTTCAAAAAAGCTGTTGAATCTGACTTACGCCTAC	75301
Db		

QY	961	AGTGGCTTTGCAAAAGTGACTGTCATTCCTATTTCTTTAAGTCCCTGAATTTATAATTTAT	1020
Db	75300	AGTGGCTTTGCAAAAGTGACTGTCATTCCTATTTCTTTAAGTCCCTGAATTTATAATTTAT	75241
QY	1021	CCTGGTTACAGCCCTTTCTGAGATGTGTGGTTTTTTTTTCCAACGTCTCTTATAGTCTGT	1080
Db	75240	CCTGGTTACAGCCCTTTCTGAGATGTGTGGTTTTTTTTTCCAACGTCTCTTATAGTCTGT	75181
QY	1081	GAATTTTCATATTTCTTTTCATACATTTTTCATGT- ---- TTTGTTTGTGTTTGTGTTGTT	1136
Db	75180	GAATTTTCATATTTCTTTTCATACATTTTTCATGTGTTTGTGTTTGTGTTTGTGTTGTT	75121
QY	1137	TTTTGGCTTTAGGTAGGCAGAACTCAGTTTCTGTGTTTTATACCCAGGAATCCTGATTGA	1196
Db	75120	TTTTGGCTTTAGGTAGGCAGAACTCAGTTTCTGTGTTTTATACCCAGGAATCCTGATTGA	75061
QY	1197	TACATCCTTCCCTTTAAAAATAAAGTATCTAAGGCTCAAGAGAGTAGGCTACCTGCCT	1256
Db	75060	TACATCCTTCCCTTTAAAAATAAAGTATCTAAGGCTCAAGAGAGTAGGCTACCTGCCT	75001
QY	1257	GAGTCTGGGAGTAAGTTAGTACCAGAGCTCGTACTAAACCCAGGTTAGCCAACTGCTTT	1316
Db	75000	GAGTCTGGGAGTAAGTTAGTACCAGAGCTCGTACTAAACCCAGGTTAGCCAACTGCTTT	74941
QY	1317	ACACAACATTTGCTCTCTCTTCAGAGTTATAGCAGTCTTGGAGAAAGAGCTACTATT	1376
Db	74940	ACACAACATTTGCTCTCTCTTCAGAGTTATAGCAGTCTTGGAGAAAGAGCTACTATT	74881
QY	1377	TTGCCAAAGACCTCAGGAGGACCAGAAACAAAGTTCTGGGATATGTGATGTAACTCTT	1436
Db	74880	TTGCCAAAGACCTCAGGAGGACCAGAAACAAAGTTCTGGGATATGTGATGTAACTCTT	74821
QY	1437	AAAAAGTTTGTGACTTCTGGCCATTAATTTGTGTATCTTAAGA CCAGATTTCAATCTTAAT	1496
Db	74820	AAAAAGTTTGTGACTTCTGGCCATTAATTTGTGTATCTTAAGA CCAGATTTCAATCTTAAT	74761
QY	1497	AGCTTAAACAACAAACAGAGATCCACAGTTTCAGCAGCTATAATAGAGTGAAATTAAGT	1556
Db	74760	AGCTTAAACAACAAACAGAGATCCACAGTTTCAGCAGCTATAATAGAGTGAAATTAAGT	74701
QY	1557	ATACAGTTGACAACATGAATATATCTCAGAAACCATGGCATCAATGAGCAAAAAAATCC	1616
Db	74700	ATACAGTTGACAACATGAATATATCTCAGAAACCATGGCATCAATGAGCAAAAAAATCC	74641
QY	1617	AGACAGAGAAGATACGTACATATGCTGCAATTTATGTGATATCTAGCAATTTGATTTGT	1676
Db	74640	AGACAGAGAAGATACGTACCATATGCTGCAATTTATGTGATATCTAGCAATTTGATTTGT	74581
QY	1677	CCAACTAGTAGCCAGTAGCCTCACATGGCTATTCAAATTTAAGTTGATTTAAATTAAGT	1736
Db	74580	CCAACTAGTAGCCAGTAGCCTCACATGGCTATTCAAATTTAAGTTGATTTAAATTAAGT	74521
QY	1737	AAGAATAAAAATTTAGTCTCTTCAGTAGCGTTAGCCACATGTAACTAGTGGCTACCAATC	1796
Db	74520	AAGAATAAAAATTTAGTCTCTTCAGTAGCGTTAGCCACATGTAACTAGTGGCTACCAATC	74461
QY	1797	AGACGGTGCAAAATATAGAAATTTTCTTTATAACAGAAAGTTCTATTTGGAAAAACAATGTT	1856
Db	74460	AGACGGTGCAAAATATAGAAATTTTCTTTATAACAGAAAGTTCTATTTGGAAAAACAATGTT	74401
QY	1857	CTAGAAAAATATACACATAAATCTATAAAAACAAAAGCAAGTCAAGTATTCCTAAGGCCA	1916
Db	74400	CTAGAAAAATATACACATAAATCTATAAAAACAAAAGCAAGTCAAGTATTCCTAAGGCCA	74341
QY	1917	GGGGTAGGGGAGATCGATTGCAAGTGGTATGAGGAAAGTTTTTGGGGTAATAGGGTTGT	1976
Db	74340	GGGGTAGGGGAGATCGATTGCAAGTGGTATGAGGAAAGTTTTTGGGGTAATAGGGTTGT	74281
QY	1977	TGGNACTTTGATTCGCATGAAGGCTACTCGGTGTCTAATGTGTCACTCTCAGACTGAA	2036
Db	74280	TGGNACTTTGATTCGCATGAAGGCTACTCGGTGTCTAATGTGTCACTCTCAGACTGAA	74221

QY	2037	CACCTTGGAAATTTGGCGAAATTTCAATTTATGTAAATTTATACCTATAAAGTAACTCTAAGAG	2096
Db	74220	CACCTTGGAAATTTGGCGAAATTTCAATTTATGTAAATTTATACCTATAAAGTAACTCTAAGAG	74161
QY	2097	GTCAAGTGTGTTGTGGAAATTTATTTTAAATCAAGTTCGAATCTTATATGAGATGATTTT	2156
Db	74160	GTCAAGTGTGTTGTGGAAATTTATTTTAAATCAAGTTCGAATCTTATATGAGATGATTTT	74101
QY	2157	TGCAAAATACATAAACAATGTTATTCATCCATTAGGTGCAATATTTTGTAGCTCCTGAAA	2216
Db	74100	TGCAAAATACATAAACAATGTTATTCATCCATTAGGTGCAATATTTTGTAGCTCCTGAAA	74041
QY	2217	ACACAGAGATGAATTAGAATAGCAAGCCTGCCCTCAAGCTGTTTCACAATCCAGTACAGGA	2276
Db	74040	ACACAGAGATGAATTAGAATAGCAAGCCTGCCCTCAAGCTGTTTCACAATCCAGTACAGGA	73981
QY	2277	GATGAGTCTATTCAAAAAATAGCTAGACTCCAGGAAGAAAGTTATAGGTGACCTTACACAA	2336
Db	73980	GATGAGTCTATTCAAAAAATAGCTAGACTCCAGGAAGAAAGTTATAGGTGACCTTACACAA	73921
QY	2337	AAAAGTCCAGATATAATTTATGTAGGACAGTAGAAGTGGGGAGGTTTCTTTATGTGGAA	2396
Db	73920	AAAAGTCCAGATATAATTTATGTAGGACAGTAGAAGTGGGGAGGTTTCTTTATGTGGAA	73861
QY	2397	AAAAGAGGAGAAATTTTGTCTTTGGAAGATGAGCAAGATGTGAATATATGCGCAGATGGA	2456
Db	73860	AAAAGAGGAGAAATTTTGTCTTTGGAAGATGAGCAAGATGTGAATATATGCGCAGATGGA	73801
QY	2457	GTTTTAAAAACATTCCTGGTGGAGGGCAGAAATATGATCCAAAGGCACAAGAGCAACCAAGAAA	2516
Db	73800	GTTTTAAAAACATTCCTGGTGGAGGGCAGAAATATGATCCAAAGGCACAAGAGCAACCAAGAAA	73741
QY	2517	ATATATCCAACTTAGAGGAAGTGCATGAAGGGGAGCAGTTGTAAATAATTTTTCATGAAT	2576
Db	73740	ATATATCCAACTTAGAGGAAGTGCATGAAGGGGAGCAGTTGTAAATAATTTTTCATGAAT	73681
QY	2577	GTAAAGTGAGAAGAAATTTGTATCATAGACACCTGAGTTTGGCAGAGTGCAATGTTCTTGGCT	2636
Db	73680	GTAAAGTGAGAAGAAATTTGTATCATAGACACCTGAGTTTGGCAGAGTGCAATGTTCTTGGCT	73621
QY	2637	CCTAGAGTCAAGAAAGAAACAAAGTGTCCCTTCTCTAAGTTATGCTCAAGTGTGCTCAAGT	2696
Db	73620	CCTAGAGTCAAGAAAGAAACAAAGTGTCCCTTCTCTAAGTTATGCTCAAGTGTGCTCAAGT	73561
QY	2697	CCAAACACCTTTCCTTCTTAAAGTACTTCTTCCCTCCCTCCATACAAATCTTAAAGTCTT	2756
Db	73560	CCAAACACCTTTCCTTCTTAAAGTACTTCTTCCCTCCCTCCATACAAATCTTAAAGTCTT	73501
QY	2757	CACAAACATCATTTAAACAGGCAGGTGATGTCAGAAAGGCAATTTCTTCTTAGACTT	2816
Db	73500	CACAAACATCATTTAAACAGGCAGGTGATGTCAGAAAGGCAATTTCTTCTTAGACTT	73441
QY	2817	CTATGTAAGTTATTTATTTCAATTTCTGCTTAAAGAGCTCTTAAAGTCTTGGAAAAAGTTT	2876
Db	73440	CTATGTAAGTTATTTATTTCAATTTCTGCTTAAAGAGCTCTTAAAGTCTTGGAAAAAGTTT	73381
QY	2877	CCACCTTGACATCAAGATATAATTTTCATGCTATTTAGTAAACCTTAGTCCCTTAAGA	2936
Db	73380	CCACCTTGACATCAAGATATAATTTTCATGCTATTTAGTAAACCTTAGTCCCTTAAGA	73321
QY	2937	GAATAAGGATGAACATAAATAAAGAAATTAATTTGGTAAATTTATATATGATTTGCCACT	2996
Db	73320	GAATAAGGATGAACATAAATAAAGAAATTAATTTGGTAAATTTATATATGATTTGCCACT	73261
QY	2997	TATTTTTCACTTGATCGTGTATGTTGCAATGCTACTGTTGTTCTCTGTTGAATTTCTAGAGAG	3056
Db	73260	TATTTTTCACTTGATCGTGTATGTTGCAATGCTACTGTTGTTCTCTGTTGAATTTCTAGAGAG	73201
QY	3057	TTTGGCTCTTTTTTCCCTGGGTCAACTCTCGCCAATTTATTTCCATAATGCAATAGAGCCAA	3116
Db	73200	TTTGGCTCTTTTTTCCCTGGGTCAACTCTCGCCAATTTATTTCCATAATGCAATAGAGCCAA	73141
QY	3117	TCTTTTTCATAATTAATTAATTTAAAAATTTGTGGCAATTTAAATTTCTGTTCTCTTAGCTT	3176

[illegible]

72061	AA	TTCA	CAGATG	TAG	CGGCTA	GTA	TAAGGA	CAGAA	TGTC	CAAGTTTTAAAT	72001
4257	AG	CTTCTC	CTTCCAT	CTCTGG	CTGA	GCAACA	AAATAAT	TTTTAT	TGAACA	CACTTT	4316
72001	AG	CTTCTC	CTTCCAT	CTCTGG	CTGA	GCAACA	AAATAAT	TTTTAT	TGAACA	CACTTT	71942
4317	GAG	TTAGAT	TTACTT	TACAGG	GAAT	TCTCT	TGAAGG	CTTTAG	NTTGTCTCA	4376	
71941	GAG	TTAGAT	TTACTT	TACAGG	GAAT	TCTCT	TGAAGG	CTTTAG	NTTGTCTCA	71882	
4377	CA	ACTTTG	CACTCT	ACTGAT	GCAC	TATTTAC	AGGTGT	CTCTGT	GACTAGG	GGGTGAA	4436
71881	CA	ACTTTG	CACTCT	ACTGAT	GCAC	TATTTAC	AGGTGT	CTCTGT	GACTAGG	GGGTGAA	71822
4437	GG	GAAGAT	GTGA	ACTCA	CCATG	TAGTGA	CCGTTA	TACAC	AGAGTGG	TTTTTTCCC	4496
71821	GG	GAAGAT	GTGA	ACTCA	CCATG	TAGTGA	CCGTTA	TACAC	AGAGTGG	TTTTTTCCC	71762
4497	CT	GTGTGG	AGTCTA	TCTTAA	CTGAG	CTTCT	GAATCAT	TATTTCA	TTCAN	TTTCCAAATCCA	4556
71761	CT	GTGTGG	AGTCTA	TCTTAA	CTGAG	CTTCT	GAATCAT	TATTTCA	TTCAN	TTTCCAAATCCA	71702
4557	CA	AAAC	CCAGTA	AGTTTAC	AGCCCAT	TATTC	AGAAAGGA	ATAAT	TATTTGTGTGTAG	4616	
71701	CA	AAAC	CCAGTA	AGTTTAC	AGCCCAT	TATTC	AGAAAGGA	ATAAT	TATTTGTGTGTAG	71642	
4617	AC	TTTCTCT	GATATTA	CACTG	ATTTGG	GAATAT	TATGAA	CAATTTAT	TGTTTCTTTCGAA	4676	
71641	AC	TTTCTCT	GATATTA	CACTG	ATTTGG	GAATAT	TATGAA	CAATTTAT	TGTTTCTTTCGAA	71582	
4677	GT	AGGTCAA	GTCAA	AGCA	AAAAAC	CAAGCA	AAAACTG	TAA	GACATAA	AGAAATAGAGT	4736
71581	GT	AGGTCAA	GTCAA	AGCA	AAAAAC	CAAGCA	AAAACTG	TAA	GACATAA	AGAAATAGAGT	71522
4737	GG	AGCGA	CTGAG	AGATTA	AAATAA	AACTAG	ATATTTT	TATTA	CAGGCAATTTGAAATA	4796	
71521	GG	AGCGA	CTGAG	AGATTA	AAATAA	AACTAG	ATATTTT	TATTA	CAGGCAATTTGAAATA	71462	
4797	AT	TTGTGC	ACTTC	CAGAA	TATTC	CAATAA	TATATTT	CCAA	TTTTTAATCTTTAAG	4856	
71461	AT	TTGTGC	ACTTC	CAGAA	TATTC	CAATAA	TATATTT	CCAA	TTTTTAATCTTTAAG	71402	
4857	AAA	ATTACT	ATATTA	TATG	TAA	GATAC	TGTGTG	AGGTAG	GATATTTAACTCA	4916	
71401	AAA	ATTACT	ATATTA	TATG	TAA	GATAC	TGTGTG	AGGTAG	GATATTTAACTCA	71342	
4917	ATA	AGGTATTTT	CTTTT	TTATTT	CGGTC	AGGCA	AGCTTCT	TAA	GGGATGTGAAAGGAT	4976	
71341	ATA	AGGTATTTT	CTTTT	TTATTT	CGGTC	AGGCA	AGCTTCT	TAA	GGGATGTGAAAGGAT	71282	
4977	AT	CTTTTCT	CTAG	CTGAG	AGAA	GATGAG	TTCTT	AA	GTAAATATAATCAAGAAATTT	5036	
71281	AT	CTTTTCT	CTAG	CTGAG	AGAA	GATGAG	TTCTT	AA	GTAAATATAATCAAGAAATTT	71222	
5037	CC	TGCTTTG	CTATTTG	ATTTG	GA	ATTTG	GA	ATTTG	GA	ATTTG	5096
71221	CC	TGCTTTG	CTATTTG	ATTTG	GA	ATTTG	GA	ATTTG	GA	ATTTG	71162
5097	AG	GGGGG	GGAGG	GA	GAATAG	TGA	AAAAA	CAAA	CAAACTTCCCTAAGCAG	5156	
71161	AG	GGGGG	GGAGG	GA	GAATAG	TGA	AAAAA	CAAA	CAAACTTCCCTAAGCAG	71102	
5157	CT	CTACAAA	ACATTTT	TAG	CGCC	CCAGAA	TAGTCA	CA	GAATCTCTCAAA	CCAGTATC	5216
71101	CT	CTACAAA	ACATTTT	TAG	CGCC	CCAGAA	TAGTCA	CA	GAATCTCTCAAA	CCAGTATC	71042
5217	CAG	ATCA	AGGA	AGTGT	TATG	CTG	GAGAGG	GTG	GACATCAT	AGCTCAGTTCAGT	5276
71041	CAG	ATCA	AGGA	AGTGT	TATG	CTG	GAGAGG	GTG	GACATCAT	AGCTCAGTTCAGT	70982
5277	TAC	AAAAGT	CCAGG	CTCTG	AAATTA	AACTCT	GATG	CCCA	TTCAT	CGCAGCATCCAA	5336
70981	TAC	AAAAGT	CCAGG	CTCTG	AAATTA	AACTCT	GATG	CCCA	TTCAT	CGCAGCATCCAA	70922

QY 5337 GACAGAGATCAGAAGTTTCAGAGATGCCTCCAGCTCCAAATTCGCAACAAAGTGTGGCT 5396
 DB 70921 GACAGAGATCAGAAGTTTCAGAGATGCCTCCAGCTCCAAATTCGCAACAAAGTGTGGCT 70862
 QY 5397 ACTATACGTCAAGGACTCTCAAGCCGTGAGAGAGGGGGAAGAAACACAGTAGAGAGGATG 5456
 DB 70861 ACTATACGTCAAGGACTCTCAAGCCGTGAGAGAGGGGGAAGAAACACAGTAGAGAGGATG 70802
 QY 5457 CCCAGCTGGTAAAGATCAGTGTTTATGAAGTTTATAGTCAATTCATGATGAATCTCATTTGGCT 5516
 DB 70801 CCCAGCTGGTAAAGATCAGTGTTTATGAAGTTTATAGTCAATTCATGATGAATCTCATTTGGCT 70742
 QY 5517 AAAATCAAGAAACGCTCCGGCTCTTTTGCARAATATGATGAAGAGAGAAAGTGCCTAAACT 5576
 DB 70741 AAAATCAAGAAACGCTCCGGCTCTTTTGCARAATATGATGAAGAGAGAAAGTGCCTAAACT 70682
 QY 5577 TCTATGTCTGATAGCAATTTGACCTTATGTCTTTTAGCTCCCGGCTTTATATCTATATAT 5636
 DB 70681 TCTATGTCTGATAGCAATTTGACCTTATGTCTTTTAGCTCCCGGCTTTATATCTATATAT 70622
 QY 5637 ACACAGTATTTGTGTATATTTTATATAATTTGTTCTCCGT 5676
 DB 70621 ACACAGTATATGTGTATATTTTATATAATTTGTTCTCCGT 70582

RESULT 7
 AC135489/c
 LOCUS AC135489 Rattus norvegicus clone CH230-436C20, WORKING DRAFT SEQUENCE, 4
 DEFINITION Rattus norvegicus clone CH230-436C20, WORKING DRAFT SEQUENCE, 4
 unorderd pieces.
 ACCESSION AC135489
 VERSION AC135489.2 GI:25139116
 KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
 SOURCE Rattus norvegicus (Norway rat)
 ORGANISM Rattus norvegicus
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 1 (bases 1 to 160574)
 Muzny,D,Marle, Metzker,M,Lees, Abramson,S, Adams,C, Alder,J.,
 Allen,C, Allen,H, Alsbrooks,S, Amin,A, Angulano,D.,
 Anyatebechi,V, Aoyagi,A, Ayodeji,M, Baca,E, Baden,H.,
 Baldwin,D, Bandaranaike,D, Barber,M, Barnstead,M, Benahmed,F.,
 Biswal,K, Blair,J, Blankenburg,K, Blyth,P, Brown,M.,
 Bryant,N, Buhay,C, Burch,P, Burrell,K, Calderon,E.,
 Cardenas,V, Carter,K, Cavazos,I, Ceasar,H, Chen,A.,
 Chacko,J, Chavez,D, Chen,G, Chen,R, Chen,Y, Chen,Z.,
 Cleveland,C, Cockrell,R, Cox,C, Coyle,M, Cree,A, D'Souza,L.,
 Davila,M,L, Davis,C, Davy-Carroll,L, De Anda,C, Dederich,D.,
 Delgado,O, Denson,S, Deramo,C, Ding,Y, Dinh,H., Divya,K.,
 Draper,H, Dugan-Rocha,S, Dunn,A, Durbin,K, Duval,B, Eaves,K.,
 Egan,A, Escotto,M, Eugene,C, Evans,C,A, Falls,T, Fan,G.,
 Fernandez,S, Finley,M, Flagg,N, Forbes,L, Foster,M, Foster,P.,
 Fraser,C,M, Gabisi,A, Ganta,R, Garcia,A, Garner,T, Garza,M.,
 Gebrgeorgis,E, Geer,K, Gill,R, Grady,M, Guerra,W, Guevara,W.,
 Gunaratne,P, Haaland,W, Hamil,C, Hamilton,C, Hamilton,K.,
 Harvey,Y, Havlak,P, Haves,A, Henderson,N, Hernandez,J.,
 Hernandez,R, Hines,S, Hladun,S,L, Hodgson,A, Hogues,M.,
 Hollins,B, Howell,S, Hulyk,S, Hume,J, Idlebird,D, Jackson,A.,
 Jackson,L, Jacob,L, Jiang,H, Johnson,B, Johnson,R, Jolivet,A.,
 Karpathy,S, Kelly,S, Kelly,S, Khan,Z, King,L, Kovar,C.,
 Kowis,C, Kraft,C,L, Lebow,H, Levan,J, Lewis,L, Li,Z, Liu,J.,
 Liu,J, Liu,W, Liu,Y, London,P, Longacre,S, Lopez,J.,
 Loresuhewa,L, Louisedge,H, Lozado,R,J, Lu,X, Ma,J.,
 Maheshwari,M, Mahindartne,M, Mahmoud,M, Malloy,K, Mangum,A.,
 Mangum,B, Mapua,P, Martin,K, Martin,R, Meenen,E.,
 Mawhney,S, McLeod,M,P, McNeill,T,Z, Meenen,E.,
 Milosavljevic,A, Miner,G, Minja,E, Montemayor,J, Moore,S.,
 Morgan,M, Morris,K, Morris,S, Munidasa,M, Murphy,M, Nair,L.,
 Nankervis,C, Neal,D, Newton,N, Nguyen,N, Norris,S.,
 Nwaokeleneh,O, Okwuonu,G, Olarnpunsagoon,A, Pal,S, Parks,K.,
 Pasternak,S, Paul,H, Perez,A, Perez,L, Pfannkoch,C.,

Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L.,
 Puazo,M., Quito,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
 Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
 Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
 Sanders,W., Savary,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
 Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D.,
 Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Soza,J.,
 Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
 Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,
 Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J.,
 Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
 Williams,G., Willson,R., Wlaczek,R., Wooden,H., Worley,K.,
 Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
 Yu,F., Zhang,J., Zhou,X., Zhou,S., Zhao,S., Dunn,D., von
 Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
 Weinstein,G. and Gibbs,R.A.
 Direct Submission
 Unpublished
 2 (bases 1 to 160574)
 Rat Genome Sequencing Consortium.
 Direct Submission
 Submitted (17-OCT-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 160574)
 Rat Genome Sequencing Consortium.
 Direct Submission
 Submitted (20-NOV-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 On Nov 20, 2002 this sequence version replaced gi:24080629.
 The sequence in this assembly is a combination of BAC based reads
 and whole genome shotgun sequencing reads assembled using Atlas
 (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
 in the feature table below represents a scaffold in the Atlas
 assembly (a 'contig-scaffold'). Within each contig-scaffold,
 individual sequence contigs are ordered and oriented, and separated
 by sized gaps filled with Ns to the estimated size. The sequence
 may extend beyond the ends of the clone and there may be sequence
 contigs within a contig-scaffold that consist entirely of whole
 genome shotgun sequence reads. Both end sequences and whole genome
 shotgun sequence only contigs will be indicated in the feature
 table.
 ----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: http://www.hgsc.bcm.tmc.edu/
 Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information
 Center project name: KBUG
 Center clone name: CH230-436C20
 ----- Summary Statistics
 Assembly program: Phrap; version 0.990329
 Consensus quality: 151333 bases at least Q40
 Consensus quality: 153444 bases at least Q30
 Consensus quality: 154259 bases at least Q20
 Estimated insert size: 155447; sum-of-contigs estimation
 Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

 * NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 4 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.
 * 1 93077: contig of 93077 bp in length
 * 93078 93177: gap of unknown length
 * 93178 113865: contig of 20688 bp in length
 * 113866 113965: gap of unknown length

* 113966 158782: contig of 44817 bp in length
 * 158783 158882: gap of unknown length
 * 158883 160574: contig of 1692 bp in length.

FEATURES

source	1. 160574	/organism="Rattus norvegicus"
		/mol_type="genomic DNA"
		/db_xref="taxon:10116"
		/clone="CH230-436C20"
misc_feature	complement(5637..6655)	
	/note="clone_boundary"	
	clone_end:117	
	site:	
	end sequence:B2203757"	
misc_feature	complement(90372..91276)	
	/note="clone_boundary"	
	clone_end:Sp6	
	site:	
	end sequence:B2203758"	
misc_feature	93178..96221	
	/note="wgs end extension"	
	clone_end:Sp6"	

ORIGIN

Query Match 11.8%; Score 669.4; DB 2; Length 160574;
 Best Local Similarity 68.5%; Pred. No. 9.1e-118;
 Matches 1192; Conservative 0; Mismatches 441; Indels 108; Gaps 15;

Qy	3981	GGCTGGTGCACAGGAATTGATTTATTTTAAATTTGAGACCTTCTATTAGTCACTA	4040
Db	65379	GGCCATAGCTACAGGACATTGATTCTCAATTTAAGAACTTTGCAATCGGGTCATCTG	65320
Qy	4041	TATTTGCTAATAGCAGGAGAAAGCCAACTCTTTAACTGCAATTAA-CAATCTATAA	4099
Db	65319	TATTTGCT-----CAGAGAGAAGTCAACACTTCAATTCGAATGAAGCAATCTATAA	65265
Qy	4100	TTAATTAGTTAAGCAATCTTCCCTTTTAACTTTTACATTTTGTGAGCAAGCTGTTGATT	4159
Db	65264	TTAATTAGTGAATGATCTTTGCTTCAAGCTTCAATTTTGAAGAAAGCTCTCTGATT	65205
Qy	4160	TGGCTGGGCTCAGCGGCTGTTTGTGAATTTCAATTTCAAGATTTAGCGCTCT	4219
Db	65204	TGGCTGGGCTCAGCGGCTGTTTGTGAATTTCAATTTCAAGATTTAGCACGCT	65145
Qy	4220	CGGCTAAGTAAGGAGAGAACTCAAGTTTAAATAGCTTCTCCCTTCCATCTCGCT	4279
Db	65144	CAGTCTCAGGAAAGGAGAGAACTCAAGTTTAAATAGCTTCTC-----CCTGGCT	65093
Qy	4280	GAAGCAACAATAAATAATTTTATGAACACATTTTGAAGTTAGTTTACTTACAGGAA	4339
Db	65092	AGAACACAATAAACAATCTTTATG-ACACATTTTGAACAGAAATTTCTTACAAGAA	65034
Qy	4340	ATGTCAAATTTCTCTGAAGGGCTTTAGATTTCTCAACATTTTGAATCTACTGATGC	4399
Db	65033	ATGTCAGATTTCT-----CTTCAGCTTGTCTCAACATTTTGAATTTCTGATGC	64983
Qy	4400	ACCTATTACAGGTGTCTCTGACATAGGGGTTGAAGGAGATTTGAATCTACCATGT	4459
Db	64982	ACTAGTTTGCAGGTGT---CTATGACTACGGCATGAATGG-AGATATGTACTTAGCATG	64927
Qy	4460	TAGTGACGGTTAGATACACAGAGTGGTTTTTTTTTCCCGCTTGTGGAGTCTATCTAACTG	4519
Db	64926	CAGGCAATTTGGGCAACGG-----CTCTATTTCCTCTTGGACTCTATTCTTAATTT	64874
Qy	4520	AGCTTCTGAATCATATTTCATTCATTAATTTCCAATTTCCAATAACAGGATTTTACGC	4579
Db	64873	AGCTTCTGAATCATATTTCATTCATTAATTTCCAATTTCCAATAACAGGATTTTACGC	64814
Qy	4580	CCATTTTCAGAAAGGAATAAATTTTGTGTGTAGACTTTT-CTGATTTACACTGAT	4638
Db	64813	CTACATTTCAGAAAGGAATAAATTTTGTGTGTAGACTTTT-CTGATTTACACTGAT	64754
Qy	4639	TTGGGAATATGAACAATTTTATGTTTCTCTTTCGAAGTAGGTCAGTCAAGCAAGCAAAAC	4698

Db	64753	TCTTCACTATATAAACAATAATGCCAGATTTTAGAGTAGAGTAAGTCAAGGAGACTAAGAG	64694
Qy	4699	CAAAAAACAGCAAACTGTAAAGACATAAAGATAGAGTGGAGCGGACTGAGAGATTAAAA	4758
Db	64693	CAACTAAAGTAGAGCTGAAGAGATAGACAGACTGTATAGAGTTTGAGGCTCACTTTAT	64634
Qy	4759	TAAACTAGAAATATTTTATTAACAGGCAATTTGAAATAAATTTGTGCATTTCAGAAATTC	4818
Db	64633	TTTTCCATTTGACGGGCTATCTGAAATAATTTCTAAATAAATATCCAAATAATAAACAAT	64574
Qy	4819	TACAATAATATATTTTCCAAATTTTAAATCTTTTAAAGAAATTTACTATATATATGTA	4878
Db	64573	TAAAGACATTTATGAAGAAACAATTAAGCTATTAATTTTAAATATCAATTTAA	64514
Qy	4879	GTACATGTGCATCTGTTTGAAGTAGGATATTTAACTCAATAAAGGTTATTTTCTTTTATT	4938
Db	64513	ATACTATACTTACATCTGTATGGATTAATAATCACCAC---AGGCTATTTTATTATT	64457
Qy	4939	CGGTCAGGCAAGCTTTCTAAGGGGATGTAAAGGGATATCTCTTTCTCT-----	4988
Db	64456	TGGAGAAGAAATATTTCTAAATACATAGGCAATCTCTCTCTCTCTCTCTCTCTCTCTCT	64397
Qy	4989	-----TAGCTGAGAGAGAGAGTCTTCTTAAG	5015
Db	64396	CT	64337
Qy	5016	TTAAATATATCAAGGAATTTCCCTGTTGCTTATTTGAGATTTGTGACCAACACAGCG	5075
Db	64336	TTAAATATATCAAGGAATTTCCCTGTTGCTTATTTGAGATTTGTGACCAACACAGCG	64277
Qy	5076	GTTGGCTGAAAGGGAACTGAAGGGGGGAGGGAGGAAATAGATGAAATAAACAACA	5135
Db	64276	GTTGGCTGAAAGGGAGGCTGAAGGGGGTTGGGGAGG-----GGAAGAGAGGGGG	64226
Qy	5136	AAACAAACTTCCCTAAGCAGCTCTACAAACAATTTTAGCCCAAGAAATAGTACAGAAA	5195
Db	64225	AAAAAATCTCTCTAAGCGCTCTGCAACAACAGTCTAGCCACAGCAGCAGTGAAGAG	64166
Qy	5196	TCCTCAATCAACACAGTATCCAGATACAGAGAGTGTATGTAGCTGGAGCAGGGTGA	5255
Db	64165	CTCTCAATCAACACAGATACAGATACAGAGAGATGTATTCAGTTGGAAACAGGGGGA	64106
Qy	5256	CACATCATCAGCTCAGTTTCAAGTTACAAAGTCCAGGCTGCTGAAATTTAACTCTGATGCCA	5315
Db	64105	CATTTATTAATCTAGT-----GACAGTCTCTGGCTCCGCTGATTTAACTCTGATGCCA	64053
Qy	5316	TTATGCGCAGATCCCAATCAACAGAGATCAGAGTTTCAGAGATGCTCCAGCTCCAAA	5375
Db	64052	TTATGCGCAGATCCCAATCAACAGAGATCAGAGTTTCAGAGATGCTCCAGCTCCAAA	63998
Qy	5376	TTGCCAACACAGTGTGGCTACTATAGCTCAAGGCTCTGAAGCCTGAGAGAGGGGA	5435
Db	63997	-----TGCCAACACAGTGTGGCTACTATAGCTCAAGGCTCTGAAGCCTGAGAGAGGGGA	63943
Qy	5436	AGAACACACAGTGTGGCTACTATAGCTCAAGGCTCTGAAGCCTGAGAGAGGGGA	5495
Db	63942	AGAACACACAGTGTGGCTACTATAGCTCAAGGCTCTGAAGCCTGAGAGAGGGGA	63883
Qy	5496	AATTTGATGAATTCATTTGGCTTAAATCAAGAAACGCTCCGCTCTTTTGCATAATGATG	5555
Db	63882	ACTTTGAAGCGCTGATTTGGCTGAAAGGAGAAATGCTCCGCTCTTAGCAAAATCTGTGTA	63823
Qy	5556	AAGGAGAGAGTGTGCTTAACTTCTATGCTGATAGCAATTTGACCTTCTTTTACGCT	5615
Db	63822	AAGGAGAGAGTGTGCTTAACTTCTATGCTGATAGCAATTTGACCTTCTTTTACGCT	63763
Qy	5616	CCGGCTTTTATATCTATATATACACAGTATTTGTGTATATTTTATATAATTTTCTCCG	5675
Db	63762	CCGGCTTTTATATCTATATATACACAGTATTTGTGTATATTTTATATAATTTTCTCCG	63703
Qy	5676	T 5676	

Query Match	11.8%;	Score	669.4;	DB 2;	Length	284005;
Best Local Similarity	68.5%;	Pred. No.	8.6e-118;			
Matches 1192;	Conservative	0;	Mismatches	441;	Indels	108; Gaps
						15;
Qy	3981	GGCTGTGGT	CACAGGAAAT	TGATTTATTTTAAATTTT	CAGAACTCTTATTTAGGT	CATCTA
Db	72189	GGCATAGCT	ACAGGACAT	TAGTATTTCTCAATTTT	AAGACTTTCCATTCGGT	CATCTG
Qy	4041	TATTTGCTAAT	TACAGGGAAG	CAAAAGCTCTTTAACTGCAAT	TAA-CAAACTAT	ATAA
Db	72129	TATTTGCT----	CAGAGAAAGT	CAAAACACTTCAATTC	CAATGAAAGCAAT	CTATAA
Qy	4100	TTAATTAGTT	TAAGCAATCTTCCCTTT	TAAGTTTACATTTGTGG	AGCAAGCTGTTGAT	T
Db	72074	TTAATTAGT	GAATGATCTTTTGCTT	CAAGCTTCACATTTTTT	TAGAAAAGCTCTCTGAT	T
Qy	4160	TGCTCGGGCT	CAGGCGGCTGTTTGCTAA	TTTTCACAAATTCACAGAT	TGTAGCGCTCT	
Db	72014	TGCTCGGGCT	CAGGCCCCCATGTGTG	AGTTTTTCAAAATTCACAGAT	TGTAGACGCT	
Qy	4220	CGGGCTAAGT	TAAAGGAGAAATG	TCAAAGTTTTTAAATAGCTTCT	CCCTCCCATCTGGCT	
Db	71954	CAGTCTGAG	AAAGGAGAAATG	TCAAGTTTTTAAATAGCTTCTC	-----CCTGGCT	
Qy	4280	GAAGCAACAA	TAAATATTTTTATGA	AAACACATTTTGAGTTAGAT	TTTACTTACAGGAA	
Db	71902	AGAACAACAA	TAAATACATCTTTATG	-AAACATTTTGAACACAGAA	TTCTTACAGAAA	
Qy	4340	ATGTCAAAT	TTTCTCGAAAGGGCTT	TAGATGTCTCACAACCTTC	GACATCTACTGATGTC	
Db	71843	ATGTCAAT	TTTCT-----CTTCAGCT	TGTCTCACAACCTTTG	CAATTTCTGATGTC	
Qy	4400	ACCTATTTA	CAGTGTCTCTGTGACT	TAGGGGTGAAGGAAGATGT	GAACTCACCATGT	
Db	71792	ACTAGTTTG	CAGGTGT--CTAGCT	ACGGCATGAATGG-AGATAT	GTACTTAGCATGT	
Qy	4460	TAGTGACGT	TAGATACACAGATGGTTTT	TTTTTCCCTGTGGAGTCTAT	CCCTAACTG	
Db	71736	CAGCGCAT	TTGGGCAACGG-----	CTCTATTTCCCTCTTGG	ACTATTTCTAAATTT	
Qy	4520	AGCTTCTGA	TATCTTCAATCTCAAT	TTCCAAATCCAAAAAC	CAGGATAAGTTTACAGC	
Db	71683	AGCTTCTGA	TATCATTTTTCAATTTCT	AAATCCACAGAACAGGAT	TGGTTTACGGC	
Qy	4580	CCATATTC	CAAAAGGAAATAAATAT	TTTTGTGTAGACTTT-CTG	ATATTACACTGAT	
Db	71623	CTACATTC	CAAAAGGAAACAAATTC	ATCATCTGAGACTTTG	CCCTGATATATTAAT	
Qy	4639	TTGGGAAT	TATGAAACAATTTTAT	TGTTTCTCTTTCGAAGT	AGTCAAGTCAAGCAAAAC	
Db	71563	TCTTCACT	ATATAAACAAATGCC	AGATTTTTTAGATGATG	TAAGTCAAGGGAGACTAAG	
Qy	4699	CAAAAA	CAGCAAAACGTGTA	AGACATAAAGATAGAGT	CGGCGGACTGAGAGAT	
Db	71503	CAACTAA	AGTGAGCTGTAAGAA	TTAGACACTGATAGAGT	TTGAGTCAACTTTAT	
Qy	4759	TAAACT	TAGAAATTTTTTAT	TAAACAGGCAATTTT	TGAAATAATTTGTG	
Db	71443	TTTTTCC	ATCTGCGGCTATCT	GAAATAATTTCTAAATA	TAAATATCCAAATAA	
Qy	4819	TACAA	TAAATATATATTTT	CCAAATTTTAAATCTTT	TAAAGAAAATTTACT	
Db	71383	TAAGA	GCATTAATAAGAA	AAACATTAACGTA	TTAACTCTTATTTTAA	


```

63836 TTAATTAGTGAATGATCTTTGGTTCAAGCTTCACATTTTATGAAAAAGCTCTCTGATT 63777
63837 TTAATTAGTGAATGATCTTTGGTTCAAGCTTCACATTTTATGAAAAAGCTCTCTGATT 63777
4160 TGGCTGGGCTCAGGCGGCTGTTGTGAATTTCAAAATTCACAGATGTTAGCGGCTCT 4219
63776 TGGCTGGGCTCAGGCGGCTGTTGTGAATTTCAAAATTCACAGATGTTAGCGGCTCT 63717
4220 CGGCTAAAGTAAAGGAGAGATGTCAGATTTTAAATAGTCTTCCCTCCATCTCTGGCT 4279
63716 CAGTCTGAGGAAGGAGAGATGTCAGATTTTAAATAGTCTCTCTCTCTCTCTCTCTCT 63665
4280 GAAGCAACAATAAATAATTTTATGAAAAACAATTTTGAAGTTAGATTTTACAGGGAA 4339
63664 AGAACACAATAAACAATCTTTATG-AAACAATTTTGAACCAAGATTTCTTACAGAAA 63606
4340 ATGTCAAATTTCTGAAAGGCTTTAGATTTGTCACAACTTTGACATCTACTGATGTC 4399
63605 ATGTCAGATTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 63555
4400 ACCTATTACAGGTGTCCTCTGACCTAGGCGGTGAAGGGAAGATGTGAACCTCACCATGT 4459
63554 ACTAGTTGAGGTGT---CTATGACTACGCGCATGATGG-AGATATGATCTTAGCATGT 63499
4460 TAGTGACGCTTAGATACACAGAGTGGTTTTTTTCCCTGTTGGAGTCTATCTCTAACTG 4519
63498 CAGCGACATTTGGGCACACGG-----CTCTATTCTCTCTCTCTCTCTCTCTCTCTCT 63446
4520 AGCTTCGATCATATTTTCAATTCATTTCCAAATCCAAACACAGATTAAGTTTACAGC 4579
63445 AGCTTCGATCATATTTTCAATTCATTTCCAAATCCAAACACAGATTAAGTTTACGGC 63386
4580 CCATATTCAGAAAGGAATAAATTTATTTTGTGTGAGACTTTT-CCTCATATTTACACTGAT 4638
63385 CTACATTCAGAAAGGAATAAATTTATTTTGTGTGAGACTTTTGTGCTGATTAATTAAT 63326
4639 TTGGGAATATATGAACAATTTTATGTTTCTTTTGAAGTGTGCTTCAAGTCAAAAGCAAAAC 4698
63325 TCTTCACTATATAAACAATGCCAGATTTTAGAGTAGATTAAGTCAAGGAGACTAAGAG 63266
4699 CAAAACAGCAAAACTGTGAAGACATTAAGATAGAGTGGAGCGACTGAGAGATTAATAA 4758
63265 CAACATAAGTGAGACGCTGAAGATTAAGACAGACTGATAGAGTTTGAGGTCAACTTTAT 63206
4759 TAAACTAGATATATTTTATTAACAGGCAATTTGAAATAATTTGTCACCTTCAGATATTC 4818
63205 TTTTCCCATTCAGCGGCTATCTGAATTAATTTCTAATAATATATCAATTAATAAACAAT 63146
4819 TACAATAATATATTTTCCAAATTTTAAATATCTTTTGAAGAAATTAATATATATATGTA 4878
63145 TAAGAACATTTATGAAGAAACAATTAACGATTAATTAACCTCTATTTTAAATATCAATTA 63086
4879 GTACATGTGATGTTTGGAGTAGGATATTTAACTCAATAAGGTTATTTCTTTTATTT 4938
63085 ATACATAAATCTTACATCTGTATGATAAATAATCAACAC---AGGCTATTTTCTTTTAA 63029
4939 CGGCTCAGGCAAACTCTTAAGGGGATGTGAAGGATATCTCTCTCTCTCTCTCTCTCTCT 4988
63028 TGGAGAGAAATAATTTCTAATAACATAGGCAATTTCTCTCTCTCTCTCTCTCTCTCTCTCT 62969
4989 -----TAGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 5015
62968 CTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 62909
5016 TTAATAATATCAAGGAATTTCCCTGCTCTGTTGATTTGAGATGTCACCAACACAGCGG 5075
62908 TTAATAATATCAAGGAATTTCCCTGCTCTGTTGATTTGAGATGTCACCAACACAGCGG 62849
5076 GTTCGCTGAAGGGAACCTGAAGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 5135
62848 GTTCGCTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 62798
5136 AAACAAACTTCCCTAAGCAGCTCTACAAAAATTTTAGCCCCAGAAATAGTACAGAAA 5195

```

```

62797 AAAAAATACTCTTAAGCGTCTCTGCCAAAAGCTCTAGCCACAGCAGCAGCAGGAG 62738
5196 TCCTCAAAATCAAAACAGTATCCAGATACAGGAAGTGTATGTAGCTGGAGAGGGTGA 5255
62737 CTCTCAAAATCAAAATCAGATACAGATACAGGAAGTGTATGTAGCTGGAGAGGGGA 62678
5256 CACTCATAGCTCAGTTCAGTTTCAAAAAGTCCAGGCTGCTGAAATTTAAATCTGTATGCCA 5315
62677 CATTTTAACTCAGT-----GACAAGTCTCTGGCTCCGCTGATTTAAATCTGTATGCCA 62625
5316 TTCTATCCAGCCTCAATCAAGCAGATCAAGATTCAGAGATGCTCCAGCTCCAAA 5375
62624 TTCTATCCAGCCTCAATCAAGCAGATCAAGATTCAGAGATGCTCCAGCTCCAAA 62570
5376 TTGCCAAACAAGTCTGGCTACTATACGTCGAAGGCTCTGAAGCCTGAGAGAGGGGA 5435
62569 -----TGCCAAACAAGTCTGGCTACTATACGTCGAAGGCTCTGAAGCCTGAGAGAGGGGA 62515
5436 AGAACACAGTACAGAGGATGCCAGCTGGTAAAGATTCGAGTGTATGAAGTTTATGTC 5495
62514 AGAACAGCTTTACAGAGGATGCCAGCTGGTAAAGATTCGAGTGTATGAAGTTTATGTC 62455
5496 AATTGATGATCTCATTTGGCTAAATCAAGAACGCTCCGCTCTTGGCAATATGATG 5555
62454 ACTTGAAGACGCTGATTTGGCTGAAGGAAGAAATGCTCCGCTCTTAGCAAAATCTGTGTA 62395
5556 AAGGAGAGAGTGCCTAAACTCTATGCTGTAGATGATTTGACCTATTTAGCTTTAGCCT 5615
62394 AAGGGGAAGTGTCTAACTTCTATGCTGTAGTGGCATTTGACCTATTTAGCTTTAGCCT 62335
5616 CCGGCTTTATATCTATATATACACAGGATTTTGTGTATATTTTATATATATTTGTTCTCG 5675
62334 CCGGCTATATATCTAGATTTCTAGTGTATATATATATATATATATATATATATATATAT 62275
5676 T 5676
62274 T 62274

```

RESULT 10

```

AC113020 214853 bp DNA linear ROD 04-OCT-2003
LOCUS Mus musculus chromosome 16, clone RP23-186N8, complete sequence.
DEFINITION AC113020
ACCESSION AC113020
VERSION AC113020.9 GI:37515087
KEYWORDS HTG.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus; Mus.
REFERENCE 1 (bases 1 to 214853)
AUTHORS Birren, B., Nusbaum, C. and Lander, E.
TITLE Unpublished
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 214853)
AUTHORS Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N.,
Anderson, S., Barna, N., Bastien, V., Boguslavsky, L., Bouckhalter, B.,
Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B.,
Choepel, S., Collange, M., Collins, S., Collymore, A., Cook, A.,
Cooke, P., DeArillano, K., Dewar, K., Diaz, J.S., Dodge, S., Fero, S.,
Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S.,
Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,
Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
Kamat, A., Karatas, A., Kells, C., LaRocque, K., Lamazares, R.,
Landers, T., Lehoczy, J., Levine, R., Liu, G., Maclean, C.,
Macdonald, P., Major, J., Marquis, N., Matthews, C., McCarthy, M.,
McEwan, P., McKernan, K., Meldrum, J., Meneus, L., Mihova, T.,
Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C.,
Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J.,
Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C.,
Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J.,
Rosetti, M., Roy, A., Santos, R., Schauer, S., Schuback, R., Seaman, S.,
Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,

```

```

Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (25-FEB-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 214853)
REFERENCE
AUTHORS
Birren,B., Nusbaum,C. and Lander,E.
Direct Submission
Submitted (05-SEP-2003) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
4 (bases 1 to 214853)
REFERENCE
AUTHORS
Anderson,M., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T.,
Boguslavskiy,L., Boukhgalter,B., Camarata,J., Chang,J., Choepel,Y.,
Collamore,A., Cook,A., Cooke,P., Corum,B., DeArellano,K.,
Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S.,
Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J., Gardyna,S.,
Graham,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B.,
Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R.,
Lindblad-Toh,K., Liu,G., Lui,X., Mabbitt,R., Maclean,C.,
Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M.,
Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J.,
Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P.,
O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,
Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P.,
Roman,J., Schauer,S., Schupbach,R., Seaman,S., Severy,P., Smith,C.,
Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M.,
Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M.,
Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X.,
Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (04-OCT-2003) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Oct 4, 2003 this sequence version replaced gi:34482079.
All repeats were identified using RepeatMasker:
Smit, A.P.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
-----
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
-----
Project Information
Center project name: L23503
Center clone name: 186_N_8
-----
FEATURES
Source
Location/Qualifiers
1..214853
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="16"
/map="16"
/clone="RP23-186N8"
/clone_lib="RPC1-23 Female Mouse BAC"
misc_feature
1..51
/notes="wgs end extension"
clone_end:SP6
misc_feature
52..57
/notes="clone boundary"
clone_end:SP6
site:EcoRI"
1042..1080
/rpt_family=" (CATATA)n"
repeat_region
1179..1367
/rpt_family="B3"
repeat_region
2436..2578
/rpt_family="L1M4c"
repeat_region
2651..2737
/rpt_family="L1M4c"

```

```

repeat_region
complement(2850..3351)
/rpt_family="RMER19B"
repeat_region
complement(3350..3516)
/rpt_family="B3"
repeat_region
3527..3891
/rpt_family="L1M4c"
repeat_region
4170..4563
/rpt_family="L1M4c"
repeat_region
4873..4952
/rpt_family="B1_MM"
repeat_region
4968..5018
/rpt_family=" (TGAA)n"
repeat_region
5056..5415
/rpt_family="MTD"
repeat_region
5554..5810
/rpt_family="L1M4c"
repeat_region
6353..6500
/rpt_family="Lx8"
repeat_region
complement(6731..7895)
/rpt_family="Lx7"
repeat_region
7896..7919
/rpt_family=" (TTTTG)n"
repeat_region
complement(7920..8304)
/rpt_family="Lx7"
repeat_region
complement(8309..9900)
/rpt_family="Lx4"
repeat_region
complement(9899..10500)
/rpt_family="Lx4"
repeat_region
10516..10639
/rpt_family="L1"
repeat_region
10634..11419
/rpt_family="L1"
repeat_region
11420..11615
/rpt_family="B2_Mm2"
repeat_region
11616..13898
/rpt_family="L1"
repeat_region
complement(13899..14142)
/rpt_family="LTRIS_MM"
repeat_region
1447..17150
/rpt_family="Lx8"
repeat_region
17251..17303
/rpt_family=" (TGG)n"
repeat_region
17523..17580
/rpt_family="AT_rich"
repeat_region
17590..17667
/rpt_family=" (GAAA)n"
repeat_region
17669..18434
/rpt_family="Lx8"
repeat_region
18558..18593
/rpt_family="AT_rich"
repeat_region
18855..19268
/rpt_family="BGLII"
repeat_region
21016..21156
/rpt_family="B1_MM"
repeat_region
21157..21199
/rpt_family=" (CAAA)n"
repeat_region
complement(21371..21767)
/rpt_family="LLME"
repeat_region
complement(21768..21878)
/rpt_family="L1"
repeat_region
21869..22524
/rpt_family="L1_MM"
repeat_region
22533..22555
/rpt_family=" (A)n"
repeat_region
complement(22556..25200)
/rpt_family="L1"
repeat_region
complement(25514..25708)
/rpt_family="B3"
repeat_region
26247..26293
/rpt_family="MT2B"
repeat_region
26294..26369
/rpt_family="B1_MM"
repeat_region
26295..26369

```


repeat_region	/rpt_family="ID_B1"	52435	TCTTTTCAATATAAAGAAAT-----GTCAGATTTTATAGTAAGTAGGTAAGAGGAGACT	52488
repeat_region	26370..26532	Qy	CAAAAAACAGCAAAACTGTAAAGACATAAAGAAATAGAGTGGAGCCGACTGAGAGATTAAAA	4758
repeat_region	26533..26553	Db		
unsure	/rpt_family="ID_B1"	52489	AAAAAAGTTAAAA---GTGAGGCATTAATAGATGAGACAGATGATAAGAGTTTGAAT	52545
repeat_region	26674..26921	Qy	TAACCTAGAAATATTTTATTAACAGGCAATTTGAAATTAATTTGTGCACCTTCAGAAATATTC	4818
repeat_region	26678..26712	Db		
repeat_region	/rpt_family="(TTTG)n"	52546	AAACTTTATTTTCCCAATTGACAGGCCATCTAAAAATAATTTCTAAAAATAATATGAAAAACA	52605
repeat_region	complement(26715..26836)	Qy	TACAATAATATATATTTTCCAAATTTTAAATATCTTTTAAAGAAAAATTTACTATATATATATGTA	4878
repeat_region	26898..26945	Db		
repeat_region	/rpt_family="(TC)n"	52606	ATAAAAAAATTAGAAAAAATTTGTAAGAAATATTAACCTCTATTTTATATATACATAAAA	52665
repeat_region	26945..26989	Qy	GTACATGTGCATGTGTTTGAGGTAGGATATTTAACTCAATAAAGGTTATTTTCTTTTATTT	4938
repeat_region	complement(26992..27264)	Db	ATCCTATTACTTATATCTG-TATGGTTAAAAAAAATCACTGCAGGCTATTTTCTTTTATTT	52724
repeat_region	30432..30458	Qy	CGGGTCAGGCAAAAGCTTCTAAAGGGGATGTGAAGGGATATCTCTCTCTCTCTCTCTCTCTCT	4988
repeat_region	/rpt_family="(T)n"	52725	TGGGACAGAAAACTATTTCTAAATATACATATACATGATCTCTCTCTCTCTCTCTCTCTCTCT	52784
repeat_region	30639..30668	Db		
repeat_region	/rpt_family="(CCCCCT)n"	4989	-----	4988
repeat_region	complement(31137..31532)	Qy	CT	52785
Query Match	11.4%; Score 649; DB 10; Length 214853;	Db		
Best Local Similarity	68.4%; Pred. No. 6.9e-114;	Qy	-TAGCTGAGAGGAGAGTGTAGTTCTAAGTTAAATAATAATCAAGGAAATTTCCCTGTCTTTG	5047
Matches 1210; Conservative	0; Mismatches 420; Indels 139; Gaps 17;	Db		
Qy	3981 GGCTGTGTCACAGGAATGATTTATTTTAAATTTTCAAGACCTTCTATTTAGTCACTCA	52845	CTCGTTGAAATGAAAGAGTGTCTAGTTTAAATATATCAAGAAATTTCCCTGTCTTTG	52904
Db	51809 GTCCATGCCCATAGAAATGATTTCTCAATTTTAAAGAACTTTTATAATCAGGTCTCTG	Qy	CTATTTGAGATTGTGACCACACAGCGGTGTGCTGAAAGGGGAAACTTGAAGGGCGGGAG	5107
Qy	4041 TATTTGCTATAGCAGGAGAGAGCAACTCTTTAACTGCAATTAACAAATCTATAAT	Db	CTATTTGAGATTGTGACCACACAGCGGTGTGCTGAAAGGGGAGGAGGAGGGGTGGG	52964
Db	51869 TATTTGCTCAGA---GGAAGAGGCAAAACAC-TTGATTTCAATGAGCAAACTATAAT	Qy	GAAGGGAATATAGATGAAAAAACAACAACAACAACTCCCTAAGCAGCTCTACAAAC	5167
Qy	4101 TAAATTAGTAAAGCAATCTCCCTTTAAAGTT--TTACATTTTGTGGACCAAGCTGTTTGT	Db	GTGGGG-----CAGGGAGAGCGCGAGAAAAAAAAGTCTCTTAAGCGTCTCTGCAAAAC	53019
Db	51924 TAAATTAGTAAAGCAATCTTTGCTCAAGTTTAAACATTTTTTTATAGCAAGCTGTCAAT	Qy	ATTTTAGCCCCAGAAATAGTCACAGAAATCTCTCAAAATCAAAACAGATATCAGATACAAGG	5227
Qy	4159 TTGCTGGGCTCAGGCGGCTGTTGTGAAATTTCAAAATTCACATTCACAGATTAGCCGCTC	Db	ATTGAGCCACACAGAACTGACAGAGCTCTCAATCAAGTCAAGATACAGATACAAGG	53079
Db	51984 TTGCTGGAGCGCAGGCGGCTGTTGTGAGTTTCAAAATTCACAGATGTTAGACACGC	Qy	AAGTGTATATAGCTGGAGCAGGCTGGACACTCATCAGCTCAGTTTCAAAAGTCC	5287
Qy	4219 TCGGCTAAGTAAAGGAGAGAAAGTCAAGTTTAAATAGCTTCTCCCTTCCATCTGGC	Db	AGATGTTATTTAGTTGGAGCAAGGGGACATTTATTAGCTCAGT-----GACAACTCC	53132
Db	52044 TCAGTCTGAGTTAGAAAGAGATGTCAAGTTTAAATAGCTTCTC-----CCTGGC	Qy	AGGCTGCTGAATTAATACTCTGATGCCATTCATGCCAGATCCNATCACGACAGATCA	5347
Qy	4279 TGAAGCAACAATAAATATTTTATGAAACACATTTTGTAGTTAGATTTACTTACAGGGA	Db	TGGCTTCTGTGATTAACCTCTGATGCCATTCATACCAAGCAACCAATCCCAAGCAAGATCA	53192
Db	52096 TGGAAACAACAATAAATCTTTATG-AAACATTTTGAAGCAGAAATTTCTTACAGGAA	Qy	GAAGTTTCAAGATGCCCTTACAAAT-----TGCCAAAGATGTGGCCACTCTACGCTCA	5407
Qy	4339 AATGTCAAATTTCTGTAAGGGCTTTAGATTGTCTCAAACTTTGACATCTACTGATGT	Db		
Db	52155 AATGTCAGATTTCT-----CTTCAGCTTGTCTCAAACTTTGACATTTTCTGATGT	Qy	AGGACTCTGAAGCGCTGAGAGAGGGGGAAGAAACAACAGTAGTAGAGAGGATGCCAGCTGGTA	5467
Qy	4399 CACCTATTTACAGGTGCTCTGTGACTAGGGGTGAGGGAAGATGTGAATCACCATG	Db	AGGCTCTAAAACCTGTGGCAGA-GAGGAAGAACAGCTTTTACAGGGGTGCCAGCTGGTA	53302
Db	52206 CACTAGTTTGCAGGTGT---CTATGGCTGTGGCATGAATGG-AGATATGTACGTAACATG	Qy	AGAACTCAGAGTGTTTATGAAGTTTTAGTCATTTAGTGAATCTCATTTGGCTAAAATCAAGAA	5527
Qy	4459 TTAGTGACCGTTAGATACACAGAGTGGTTTCTTTTCCCTCTGTTGGAGTCTATCTAACT	Db	AGAACTGACCGTTTATGATGCTCTGGTTACTTGAAGACTCTCATTTGGCTGAAAGGAGAA	53362
Db	52262 TCAGTGACACTTGGGCACACAG-----CTCTAGTTTCCCTCTTGGACTCTATTTCTAAT	Qy	ACGCTCCGCTCTTTGCAAAATATGATGAAGGAGAGAGTGCCTTAAACTTCTATGCTCTGA	5587
Qy	4519 GAGCTTCTGAATCATATTTCAATTTCAAAATCCAAACCAAGGATAGTTTACAG	Db	ACGCCCCCTCTTTGCAAAATCTGAGTAAAGGGGGAAGTGTCTTAAACTTCTATGCTCTGA	53422
Db	52315 TAGCTTCTGAATCATATTTCAATTTCAATTTTCAATTTCAATTTCAATTTTCAATTTTCA	Qy	TAGCAATTTGACCTTATTTAGCTTCCCGCTTTTATAGCTTCTCTCTCTCTCTCTCTCTCTCTCT	5647
Qy	4579 CCCATATTTCAAGAGGAATAAATTTATTTTGTGTGATAGACTTTCTCTGATATATCACTGAT	Db	TGGCAATTTGACCTTATTTGCTTTTTCAGCTTCTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCT	53482
Db	52375 CCTACATTTCAAGAGGAACAAATTTTATTTATCATGAGACTTTTGCCTGATATATATAAT	Qy	TGTGTATATTTTATATATTTTGTCTCCGT	5676
Qy	4639 TTGGGAATATATGACAAATTTTATGTTTCTCTTTCGAAGTAGGTCAAGTCAAGCAAGCAAAAC	Db	TATGTATATTTTATAGAAATTTGCTTCCCAT	53511

```

RESULT 11
AC091240/c
LOCUS          56268 bp      DNA      linear      HTG 07-APR-2001
DEFINITION    Mus musculus clone RP23-16N1, LOW-PASS SEQUENCE SAMPLING.
ACCESSION     AC091240
VERSION       AC091240.1  GI:13560417
KEYWORDS      HTG; HTGS PHASE0.
SOURCE        Mus musculus (house mouse)
ORGANISM      Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE     1 (bases 1 to 56268)
AUTHORS       Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE         Mus musculus, clone RP23-16N1
JOURNAL
AUTHORS
REFERENCE     2 (bases 1 to 56268)
AUTHORS       Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,
Barna,N., Bastien,V., Boguslavskiy,L., Boukhgalter,B., Brown,A.,
Camarata,J., Campopiano,A., Chang,J., Choepel,Y., Colangelo,M.,
Collins,S., Collymore,A., Cooke,P., Dearellano,K., Dewar,K.,
Diaz,J.S., Dodge,S., Faro,S., Ferreira,P., Fitzhugh,W., Gage,D.,
Galan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
Grand-Pierre,N., Hagos,B., Heaford,A., Horton,L., Hulme,W.,
Iliev,I., Johnson,R., Jones,C., Karatas,A., LaRocque,K.,
Lamazares,R., Landers,T., Lehotzky,J., Levine,R., Liu,G.,
MacLean,C., Macdonald,P., Marquis,N., Matthews,C., McCarthy,M.,
McEwan,P., McKernan,K., McPheters,R., Meldrum,J., Meneus,D.,
Mihova,T., Mienga,V., Murphy,T., Naylor,J., Nguyen,C., Norbu,C.,
Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J.,
Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C.,
Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J.,
Rosetti,M., Roy,A., Santos,R., Schauer,S., Schuback,R., Seaman,S.,
Severy,P., Sougniez,C., Spencer,B., Stange-Thomann,N.,
Stojanovic,N., Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S.,
Theodore,J., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (07-APR-2001) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
All repeats were identified using RepeatMasker:
Smith, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L13295
Center clone name: 16_N_1
-----
* NOTE: This record contains 70 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.
*
* 1 713: contig of 713 bp in length
* 377: gap of 100 bp
* 378 813: contig of 704 bp in length
* 378 1517: gap of 100 bp
* 459 1618: contig of 697 bp in length
* 459 2314: gap of 100 bp
* 459 2415: contig of 677 bp in length
* 459 3092 3191: gap of 100 bp
*
* 3192 3877: contig of 686 bp in length
* 3877: gap of 100 bp
* 3978 4658: contig of 681 bp in length
* 4559 4758: gap of 100 bp
* 4759 5468: contig of 710 bp in length
* 5469 5569: gap of 100 bp
* 6273: contig of 705 bp in length
* 6374 6794: gap of 100 bp
* 6794 7094: contig of 721 bp in length
* 7095 7194: gap of 100 bp
* 7195 7890: contig of 696 bp in length
* 7891 7990: gap of 100 bp
* 7991 8716: contig of 726 bp in length
* 8717 8816: gap of 100 bp
* 8817 9527: contig of 711 bp in length
* 9528 9627: gap of 100 bp
* 9628 10338: contig of 711 bp in length
* 10339 10438: gap of 100 bp
* 10439 11128: contig of 690 bp in length
* 11129 11228: gap of 100 bp
* 11229 11949: contig of 721 bp in length
* 11950 12049: gap of 100 bp
* 12050 12758: contig of 709 bp in length
* 12759 12858: gap of 100 bp
* 12859 13572: contig of 714 bp in length
* 13573 13672: gap of 100 bp
* 13673 14357: contig of 685 bp in length
* 14358 14457: gap of 100 bp
* 14458 15167: contig of 710 bp in length
* 15168 15267: gap of 100 bp
* 15268 15949: contig of 682 bp in length
* 15950 16049: gap of 100 bp
* 16050 16738: contig of 689 bp in length
* 16739 16838: gap of 100 bp
* 16839 17543: contig of 707 bp in length
* 17544 17645: gap of 100 bp
* 17646 18372: contig of 727 bp in length
* 18373 18472: gap of 100 bp
* 18473 19206: contig of 734 bp in length
* 19207 19306: gap of 100 bp
* 19307 20017: contig of 711 bp in length
* 20018 20117: gap of 100 bp
* 20118 20835: contig of 718 bp in length
* 20836 20935: gap of 100 bp
* 20936 21631: contig of 696 bp in length
* 21632 21731: gap of 100 bp
* 21732 22419: contig of 688 bp in length
* 22420 22519: gap of 100 bp
* 22520 23215: contig of 696 bp in length
* 23216 23315: gap of 100 bp
* 23316 24016: contig of 701 bp in length
* 24017 24116: gap of 100 bp
* 24117 24826: contig of 710 bp in length
* 24827 24926: gap of 100 bp
* 24927 25637: contig of 711 bp in length
* 25638 25737: gap of 100 bp
* 25738 26460: contig of 723 bp in length
* 26461 26560: gap of 100 bp
* 26561 27261: contig of 701 bp in length
* 27262 27361: gap of 100 bp
* 27362 28075: contig of 714 bp in length
* 28076 28175: gap of 100 bp
* 28176 28880: contig of 705 bp in length
* 28881 28980: gap of 100 bp
* 28981 29785: contig of 705 bp in length
* 29786 30479: contig of 694 bp in length
* 30480 30579: gap of 100 bp
* 30580 31254: contig of 675 bp in length
* 31255 31354: gap of 100 bp
* 31355 32042: contig of 688 bp in length
* 32043 32142: gap of 100 bp
* 32143 32847: contig of 705 bp in length

```

27246	Db	GTCTTTTGCTATTGTGAGTTGTGACCAACAAGCGCGTGTGGCTGAAAGGGAGCGAGAAGGG	27187
5101	Qy	CGGGGAGGGAGGAGAAATAGATGAAAAAAA CAAAA CAAAA CAAAACTTCCTTAAGCAGTCT	5160
27186	Db	GGTGGGGTGGGG-----CAGGGAGAGCCGAGAAAAAAAAGATCTCTTAAGCGTCTCT	27132
5161	Qy	ACAAACATTTTAGCCCCAGAAATAGTCA CAGAAATCTCAAA TCAACACAGTATCCAGA	5220
27131	Db	GCAAAACATTTGTAGCCACAGCAGAACTCACAGGAGCTCTCAAA TCAAGTCAGATACAGA	27072
5221	Qy	TACAAGGAAGTCTTATGTAGCTGGAGCAGGGGTGACACATCATCAGCTCAGTTCAGTTTACA	5280
27071	Db	TACNAGGAGATGTTATT CAGTTGGAGCAAGGGGACATTTATTAGTTCAGT-----GA	27019
5281	Qy	AAAGTCGAGGCTGCTGAAATTTAAACTCTGTATGCCATTTCATGCCAGCATCCAA TCAACGACA	5340
27018	Db	CAAGTCCTGGCTTCTGTGATTAAACTCTGTATGCCATTTCATACCAAGCAGCCCAATCCCAAGC	26959
5341	Qy	GAGATCAGAAGTTCAGAGATGCTCCAGCTCCAAATTCGCAACCAACAACTGTGGCTACTA	5400
26958	Db	AGATCAGNAGTTTCAAGATGCTTACAAT-----TGCCACAAGTGTGGCCACTC	26908
5401	Qy	TACGTCAAGGACTCTGAAGCCGTGAGAGAGGGGGGAAGAACAA CAGTAGAGAGGATGCCCA	5460
26907	Db	TACGTCAAGGGCTCTAAAACTGTGCAGAGAGGGAAGAACAGCTTTTACAGGGGGTGCACA	26849
5461	Qy	GCTGGTGAAGAAATCGAGTGTATTATGAAGTTTATAGTCAATTTGATGAATCTCATTTGGCTAAAA	5520
26848	Db	GCTGGTGAAGAAATGACGTTTATGATGCTCTGGTTACTTTGAAGACTCTCATTTGGCTGAAA	26789
5521	Qy	TCAAGAAACGCTCCGCTCTTTTGC AAAATATGATGAAGGAGAGAGTGCCTAAACTTCTA	5580
26788	Db	GGAAGAAACGCCCGCTCTTTTGCAAACTCTGAGTAAAGGGGGAAGTGTCTAAACTTCTA	26729
5581	Qy	TGCTGATAGCATTTGACCCATTGCTTTTAGCTCTCCGGCTTTATATCTATATATACAC	5640
26728	Db	TGTCGTATGGCAATTTGACCCATTGCTTTGACGCTCTCGGCTACATACCTAGATATTCTC	26669
5641	Qy	AGGTATTTCTGTATATTTTATATAATTTGTTCTCCCT	5676
26668	Db	AGGTGATATGATATATTTTATAGAATTTGCTTCCCAT	26633

RESULT 12

HSP63G01

LOCUS

DEFINITION

Homo. sapiens p63 protein (p63) gene, exon 1.

HSP63G01

835 bp

DNA

linear

PRI 04-JAN-2001

AF124528
AF124528.1 GI:12024731
1 of 13
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 835)
Yang, A., Kaghad, M., Wang, Y., Gillett, E., Fleming, M.D., Dotsch, V.,
Andrews, N.C., Caput, D. and McKeon, F.
p63, a p53 homolog at 3q27-29, encodes multiple products with
transactivating, death-inducing, and dominant-negative activities
Mol. Cell 2 (3), 305-316 (1998)
98448095
MEDLINE
PUBMED
9774969
2 (bases 1 to 835)
Hagiwara, K., McMenamin, M.G. and Harris, C.C.
Direct Submission
Submitted (29-JAN-1999) Laboratory of Human Carcinogenesis,
National Cancer Institute, Building 37, Room 2C22, 37 Convent
Drive, Bethesda, MD 20892, USA
Location/Qualifiers
1..835
/organism="Homo sapiens"
/mol type="genomic DNA"

```
exon          /db_xref="taxon:9606"
              216..305
              /gene="p63"
              /number=1

ORIGIN
Query Match      3.8%; Score 215; DB 9; Length 835;
Best Local Similarity 100.0%; Pred. No. 8e-31;
Matches 215; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5462 CTGTAAGAAATCGAGTGTATGAAGTTTAGTCAATTGATGAATCTCATTCATGCTAAAT 5521
Db 1 CTGTAAGAAATCGAGTGTATGAAGTTTAGTCAATTGATGAATCTCATTCATGCTAAAT 60

Qy 5522 CAAGAAACGCTCCGCTCTTTGCAAAATATGTAAGGAGAGAGTGCCTAAACTTCTAT 5581
Db 61 CAAGAAACGCTCCGCTCTTTGCAAAATATGTAAGGAGAGAGTGCCTAAACTTCTAT 120

Qy 5582 GTCTGATAGCAATTTGACCCCTATTGCTTTTAGCCTCCCGGCTTTATATCTATATACACA 5641
Db 121 GTCTGATAGCAATTTGACCCCTATTGCTTTTAGCCTCCCGGCTTTATATCTATATACACA 180

Qy 5642 GGTATTGTGTATATTTATATATATTTGTTCTTCCT 5676
Db 181 GGTATTGTGTATATTTATATATATTTGTTCTTCCT 215

RESULT 13
AC103882/c    159249 bp DNA linear PRI 15-OCT-2002
LOCUS        Homo sapiens BAC clone RP11-733G6 from 2, complete sequence.
DEFINITION   AC103882
ACCESSION    AC103882
VERSION      AC103882.5 GI:23130727
KEYWORDS     HTG.
SOURCE       Homo sapiens (human)
ORGANISM     Homo sapiens
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
              Sulston, J.E. and Waterston, R.
              1 (bases 1 to 159249)
              Toward a complete human genome sequence
              Genome Res. 8 (11), 1097-1108 (1998)
              99063792
              PUBMED 9847074
              2 (bases 1 to 159249)
              Wang, C. and Bielicki, L.
              The sequence of Homo sapiens BAC clone RP11-733G6
              Unpublished (2001)
              3 (bases 1 to 159249)
              Waterston, R.H.
              Direct Submission
              Submitted (29-NOV-2001) Genome Sequencing Center, Washington
              University School of Medicine, 4444 Forest Park Parkway, St. Louis,
              MO 63108, USA
              4 (bases 1 to 159249)
              Waterston, R.H.
              Direct Submission
              Submitted (30-MAR-2002) Genome Sequencing Center, Washington
              University School of Medicine, 4444 Forest Park Parkway, St. Louis,
              MO 63108, USA
              5 (bases 1 to 159249)
              Waterston, R.H.
              Direct Submission
              Submitted (18-SEP-2002) Genome Sequencing Center, Washington
              University School of Medicine, 4444 Forest Park Parkway, St. Louis,
              MO 63108, USA
              6 (bases 1 to 159249)
              Waterston, R.
              Direct Submission
              Submitted (03-OCT-2002) Department of Genetics, Washington
              University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
              7 (bases 1 to 159249)
              Waterston, R.
```

```
TITLE
JOURNAL
COMMENT
Direct Submission
Submitted (15-OCT-2002) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On Sep 18, 2002 this sequence version replaced gi:19703341.
----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc
Contact: sapiens@wustl.edu
----- Summary Statistics
Center project name: H_NH0733G06
-----
```

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:

The RPCI-11 human BAC library was made from the blood of one male donor, as described by Osogawa, K., Woon, P.Y., Zhao, B., Frengen, E., Tatenio, M., Catanese, J. and de Jong, P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at <http://www.chori.org>
VECTOR: pBACE3.6

NEIGHBORING SEQUENCE INFORMATION:

This sequence is not the entire insert of the clone. This clone is overlapped by AC131754 and AC096656.

There is a simple sequence repeat from base 46448 to 46739. This region does not meet required finishing standards.

```
FEATURES
source
1..159249
/organism="Homo sapiens"
/mol_type="Genomic DNA"
/db_xref="taxon:9606"
/chromosomes="2"
/map="2"
/clone="RP11-733G6"
/clone_lib="RPCI-11"
repeat_region 102..178
repeat_region 179..571
repeat_region 572..626
repeat_region 765..942
repeat_region 1024..1916
repeat_region 1917..2217
repeat_region 2218..2499
repeat_region 2499..3481
```

```
repeat_region 3482..3782 /rpt_family="Alu"
repeat_region 3783..3929 /rpt_family="L1"
repeat_region 3930..4251 /rpt_family="L1"
repeat_region 4252..4764 /rpt_family="Alu"
repeat_region 4776..4986 /rpt_family="L1"
repeat_region 4994..5023 /rpt_family="L1"
repeat_region 5024..5176 /rpt_family="(TA)n"
repeat_region 5236..5404 /rpt_family="L1"
repeat_region 5520..5705 /rpt_family="L1"
repeat_region 5747..5777 /rpt_family="L1"
repeat_region 5782..5972 /rpt_family="AT_rich"
repeat_region 6450..6474 /rpt_family="L1"
repeat_region 6998..7244 /rpt_family="(TG)n"
repeat_region 7948..8299 /rpt_family="MIR"
repeat_region 8868..9128 /rpt_family="MER2_type"
repeat_region 9299..9510 /rpt_family="MER2_type"
repeat_region 9516..9711 /rpt_family="ERVL"
repeat_region 9717..10011 /rpt_family="ERVL"
repeat_region 10275..10313 /rpt_family="Alu"
repeat_region 10375..10528 /rpt_family="AT_rich"
repeat_region 10529..10659 /rpt_family="(TA)n"
repeat_region 10696..11412 /rpt_family="(TATAA)n"
repeat_region 11440..11544 /rpt_family="ERVL"
repeat_region 11569..11927 /rpt_family="ERVL"
repeat_region 12292..12495 /rpt_family="MaLR"
repeat_region 12856..13041 /rpt_family="ERVL"
repeat_region 13197..14809 /rpt_family="ERVL"
repeat_region 14829..14876 /rpt_family="L1"
repeat_region 14936..15187 /rpt_family="MIR"
repeat_region 15243..15523 /rpt_family="MER1_type"
repeat_region 15728..15999 /rpt_family="Alu"
repeat_region 16186..16418 /rpt_family="Alu"
repeat_region 16727..17022 /rpt_family="MaLR"
repeat_region 17166..17468 /rpt_family="Alu"
repeat_region 17670..17868 /rpt_family="Alu"
repeat_region 18530..18674 /rpt_family="MIR"
repeat_region 19048..19069 /rpt_family="MIR"

/rpt_family="AT_rich"
20886..21207
/rpt_family="CRI"
21379..22209
/rpt_family="L1"
22342..22558
/rpt_family="MIR"
22572..22872
/rpt_family="CRI"
23219..23264

Query Match 3.7%; Score 212.4; DB 9; Length 159249;
Best Local Similarity 63.3%; Pred. No. 1.6e-30;
Matches 413; Conservative 0; Mismatches 221; Indels 18; Gaps 5;

QY 3293 TGGTGGCTCATGCTGTAATCCCAACACTTTGGGAAGCCGGGTCGGGAGGATAGCTTGAG 3352
DB 85766 TGGTGGCTTCGCTGAACTCCCAAGCATTGGGAGGCCAAGGTGGGAGGATCTTTGAG 85707

QY 3353 TCCAGCAGTTTGAGACCAGTCAGGGCAACACAGCAAGACCCCATATCTAAAAAACAATA 3412
DB 85706 CTCAGGCAATTTGAGACCAGCTGGGCAATATAAAAAAGGTCTCGTGTCTACTAAAAATTACA 85647

QY 3413 CAAAAAATAATACCTGGGTATGTTGTCTCACTGTAGTCCCAAGTCACACAGGAGCT 3472
DB 85646 -----AAAATTAGCTGGGTGTGGTGATCATGCTGTAGTCCAGCTACTCAGGAGGCT 85593

QY 3473 GAGCGAAGGATCACTTTGAGCCAGGAGGTTCAGGCTGCAGTGATCCATGAACGCCGTG 3532
DB 85592 GAAGCGAAGGATCACTTCGAGCCAGGAGGTAGAGGTGCAGTGAGCCAGATGGCACCA 85533

QY 3533 CTACACT-CAGTCTGGGTGACAGTGCAGGAAGCTGTCTCAAAAAATAATAATAATAAAAA 3591
DB 85532 CTGTACTCCAGCCTGGGTGACAGCAAGACTCTGTCTCAAAAAATAATAATAATAAAAAAG 85473

QY 3592 ATAACTTTAAAAACAATAATAATAATAATAATAATAATAATAATAATAATAATAATAA 3649
DB 85472 AAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAG 85413

QY 3650 TTCAATATTCATTATTTGGGAGCTATATCCCTGGGAGTTAATTTAAAAATAATTTAGAAGA 3709
DB 85412 AAGCAATTTGTCCTGGGAAAAATCACTAATATCCAGTGTCTGTGTATCACATAGCAGA 85353

QY 3710 GTTCTTCTCATTTCTCAGAGAGCTCGAATTGTAAATATCAGAGCTAGAGGACACTAG 3769
DB 85352 AAACAGCCCTTCAGTTAGCGAACTGGGTTGAGGAGAGCAAGGCAAGCAAGAA 85293

QY 3770 GGCTGGCCACTCCAAAGTGTGTCAGGACCCAGCAGCATCAAGTAACTTGGGAACTGT 3829
DB 85292 GGTGCTAGGGAAGAGAACTGGAGAAATGTTGCAGGCAC-----TGCCTGGGAGCTTGT 85240

QY 3830 TAGAATGTCAGAGTCTTAGGCTCACCCAGAGCTACTGAACCA--CAATCTGCATTAAAC 3887
DB 85239 CAGAAATGTCAGAGTCTTTGGGCTTCTCCCAAGACCCACTGAATCAGGCACTGATTTTAAAC 85180

QY 3888 AAGATTTCTAGGTGCTCTCAGGGCACAATAAAACTTCAGAGCTCTGCACCTA 3939
DB 85179 AAGACCTCTAGGTGATTCCTGCACAGGTTAAACTTCAGAGCACTGCACAA 85128

RESULT 14
AC087334 60323 bp DNA linear HTG 30-DEC-2001
LOCUS Homo sapiens chromosome 17 clone RP11-73H6 map 17, LOW-PASS
DEFINITION AC087334 SEQUENCE SAMPLING.
ACCESSION AC087334
VERSION AC087334.3 GI:17998705
KEYWORDS HTG; HTGS PHASE0.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 60323)
```

AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS

Birren,B., Linton,L., Nusbaum,C. and Lander,E.
Homo sapiens chromosome 17, clone RP11-73H6
Unpublished
2 (bases 1 to 60323)
Barren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,
Barra,N., Bastien,L., Boguslavsky,L., Boukhgalter,B., Brown,A.,
Canarata,J., Campopiano,A., Choepel,Y., Colangelo,M., Collins,S.,
Collimore,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S.,
Dodge,S., Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galsgan,J.,
Gardyna,S., Ginde,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
Jones,C., Karatas,A., LaRocque,K., Lamazares,R., Landers,T.,
Lehoczy,J., Levine,R., Liu,G., MacLean,C., Macdonald,P.,
Marquis,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K.,
McPheters,R., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
Murphy,T., Naylor,J., Nguyen,C., Norbu,C., Norman,C.H.,
O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,
Phunkhang,P., Pierre,N., Pollara,V., Raymond,C., Retta,R.,
Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M.,
Roy,A., Santos,R., Schauer,S., Schuback,R., Seaman,S., Severy,P.,
Sougnez,C., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Travers,M., Travis,N., Trigilio,J., Vassiliev,H., Viel,R., Vo,A.,
Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J.,
Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (28-DEC-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Dec 30, 2001 this sequence version replaced gi:12658045.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

TITLE
JOURNAL
COMMENT

Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence.submissions@genome.wi.mit.edu
----- Project Information
Center project name: L10827
Center clone name: 73_H_6

* NOTE: This record contains 72 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.

* 1
* 722: contig of 722 bp in length
* 723
* 822: gap of 100 bp
* 823
* 1575: contig of 753 bp in length
* 1576
* 1675: gap of 100 bp
* 1676
* 2434: contig of 759 bp in length
* 2435
* 2534: gap of 100 bp
* 2535
* 3262: contig of 728 bp in length
* 3263
* 3362: gap of 100 bp
* 3363
* 4097: contig of 735 bp in length
* 4098
* 4197: gap of 100 bp
* 4198
* 4951: contig of 754 bp in length
* 4952
* 5051: gap of 100 bp
* 5052
* 5810: contig of 759 bp in length
* 5811
* 5910: gap of 100 bp
* 5911
* 6680: contig of 770 bp in length
* 6681
* 6780: gap of 100 bp
* 6781
* 7511: contig of 731 bp in length
* 7512
* 7611: gap of 100 bp
* 7612
* 8341: contig of 730 bp in length
* 8342
* 8441: gap of 100 bp
* 8442
* 9174: contig of 733 bp in length
* 9175
* 9274: gap of 100 bp
* 9275
* 10016: contig of 742 bp in length
* 10017
* 10116: gap of 100 bp
* 10117
* 10860: contig of 744 bp in length
* 10861
* 10960: gap of 100 bp
* 10961
* 11709: contig of 749 bp in length
* 11710
* 11809: gap of 100 bp
* 11810
* 12383: contig of 574 bp in length
* 12384
* 12483: gap of 100 bp
* 12484
* 13244: contig of 761 bp in length
* 13245
* 13344: gap of 100 bp
* 13345
* 14041: contig of 697 bp in length
* 14042
* 14141: gap of 100 bp
* 14142
* 14894: contig of 753 bp in length
* 14895
* 14994: gap of 100 bp
* 14995
* 15751: contig of 757 bp in length
* 15752
* 15851: gap of 100 bp
* 15852
* 16575: contig of 724 bp in length
* 16576
* 16675: gap of 100 bp
* 16676
* 17401: contig of 726 bp in length
* 17402
* 17501: gap of 100 bp
* 17502
* 18237: contig of 736 bp in length
* 18238
* 18337: gap of 100 bp
* 18338
* 19073: contig of 736 bp in length
* 19074
* 19173: gap of 100 bp
* 19174
* 19907: contig of 734 bp in length
* 19908
* 20007: gap of 100 bp
* 20008
* 20762: contig of 755 bp in length
* 20763
* 20862: gap of 100 bp
* 20863
* 21569: contig of 707 bp in length
* 21570
* 21669: gap of 100 bp
* 21670
* 22380: contig of 711 bp in length
* 22381
* 22480: gap of 100 bp
* 22481
* 23231: contig of 751 bp in length
* 23232
* 23331: gap of 100 bp
* 23332
* 24101: contig of 770 bp in length
* 24102
* 24201: gap of 100 bp
* 24202
* 24959: contig of 758 bp in length
* 24960
* 25060: gap of 100 bp
* 25061
* 25780: contig of 721 bp in length
* 25781
* 25880: gap of 100 bp
* 25881
* 26614: contig of 734 bp in length
* 26615
* 26714: gap of 100 bp
* 26715
* 27436: contig of 722 bp in length
* 27437
* 27536: gap of 100 bp
* 27537
* 28277: contig of 741 bp in length
* 28278
* 28377: gap of 100 bp
* 28378
* 29108: contig of 731 bp in length
* 29109
* 29208: gap of 100 bp
* 29209
* 29663: contig of 755 bp in length
* 29664
* 30063: gap of 100 bp
* 30064
* 30831: contig of 768 bp in length
* 30832
* 30931: gap of 100 bp
* 30932
* 31641: contig of 710 bp in length
* 31642
* 31741: gap of 100 bp
* 31742
* 32490: contig of 749 bp in length
* 32491
* 32590: gap of 100 bp
* 32591
* 33338: contig of 748 bp in length
* 33339
* 34170: contig of 732 bp in length
* 34171
* 34270: gap of 100 bp
* 34271
* 35010: contig of 740 bp in length
* 35011
* 35110: gap of 100 bp
* 35111
* 35936: contig of 826 bp in length
* 35937
* 36036: gap of 100 bp
* 36037
* 36800: contig of 764 bp in length
* 36801
* 36900: gap of 100 bp
* 36901
* 37632: contig of 732 bp in length
* 37633
* 37732: gap of 100 bp
* 37733
* 38436: contig of 704 bp in length
* 38437
* 38536: gap of 100 bp
* 38537
* 39309: contig of 773 bp in length
* 39310
* 39409: gap of 100 bp


```
* 39410 40153: contig of 744 bp in length
* 40154 40253: gap of 100 bp
* 40254 40981: contig of 728 bp in length
* 40982 41081: gap of 100 bp
* 41082 41819: contig of 738 bp in length
* 41820 41919: gap of 100 bp
* 41920 42646: contig of 727 bp in length
* 42647 42747: gap of 100 bp
* 42748 43477: contig of 731 bp in length
* 43478 43578: gap of 100 bp
* 43579 44310: contig of 733 bp in length
* 44311 44410: gap of 100 bp
* 44411 45176: contig of 766 bp in length
* 45177 45276: gap of 100 bp
* 45277 46043: contig of 767 bp in length
* 46044 46143: gap of 100 bp
* 46144 46849: contig of 706 bp in length
* 46850 46949: gap of 100 bp
* 46950 47701: contig of 752 bp in length
* 47702 47801: gap of 100 bp
* 47802 48543: contig of 742 bp in length
* 48544 48644: gap of 100 bp
* 48644 49411: contig of 768 bp in length
* 49412 49511: gap of 100 bp
* 49512 50236: contig of 725 bp in length
* 50237 50336: gap of 100 bp
* 50337 51067: contig of 731 bp in length
* 51068 51167: gap of 100 bp
* 51168 51895: contig of 728 bp in length
* 51896 51995: gap of 100 bp
* 51996 52703: contig of 708 bp in length
* 52704 52803: gap of 100 bp
* 52804 53550: contig of 747 bp in length
* 53551 53651: gap of 100 bp
* 53651 54364: contig of 714 bp in length
* 54365 54464: gap of 100 bp
* 54465 55227: contig of 763 bp in length
* 55228 55327: gap of 100 bp
* 55328 56118: contig of 791 bp in length
* 56119 56219: gap of 100 bp
* 56219 57061: contig of 743 bp in length
* 57062 57773: gap of 100 bp
* 57773: contig of 712 bp in length

Query Match      3.6%; Score 205; DB 2; Length 60323;
Best Local Similarity 57.2%; Pred. No. 4.4e-210;
Matches 282; Conservative 0; Mismatches 1; Gaps 1;

Qy 3264 TTTAAATTTTAAATAAATAGATGAGTGTGGTGCCTCATGCTGTGTAATCCCAACACTTT 3323
Db TCTAATTTTAAGAAAAATGCTTGGGATGTGGTGCCTCATGCTGTGTAATCCTAACACTTT 9694

Qy 3324 GGGAGCGGGTCCGGAGGATAGCTTGTAGTCCAGAGTTTGAGACCGAGTCAGGCGCAACAC 3393
Db GGGAGCGCAAGCGGATGATGCTGCTGAGCTCAGCGCTTTCGCAACCGAGCTGGCGCAACAC 9754

Qy 3384 AGCAGAGCCCATCTTAAATAAACAACAACAACAATAATACCTGGTATGTTGCT 3443
Db GGTGAACCCCGCTCTACTTAAATAAATAACAATAAATAATACCTGGCATGGTATG 9814

Qy 3444 CACCTGTAGTCCAGCTACAGAGAGCTGAGGCGAGGAGTCACTTGAAGCCAGGAGGT 3503
Db CACCTGTATCCAGCGACTCAAGAGGCTGAGGCGAGGAGTCACTTGAACCCAGGAGGC 9874

Qy 3504 TGAGGCTGCAGTGATCCATGAACCGCTGCTACACT-CAGTCTGGGTGACAGTCCAGAA 3562
Db GGAGTGCAGTGAGCCAAAGATCGGCCACTGCCTACAGCTGGCGCAGACGAGAG 9934

Qy 3563 GCTGTCTCAAAATAATTAATAATAAATAAATACTTTTAAATAAACAATAATTAATAAT 3622
Db TCCGTCTCAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAAT 9994

Qy 3623 TTTAAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACA 3682
```

```
Db 9995 TAATTAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 10054
Qy 3683 GGAAGTAAATTTAAATAATTTAGAGAGTTCCTCTCATTTCTCTAGAGAGTCCGAATGT 3742
Db 10055 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 10114

Qy 3743 AAATATCAGAGT 3755
Db 10115 NNAGCGAGATCT 10127

RESULT 15
LOCUS AL158198
DEFINITION Human DNA sequence from clone RP11-239H6 on chromosome 6 Contains an STS, GSSs and CpG islands, complete sequence.
ACCESSION AL158198
VERSION AL158198.14 GI:9944140
KEYWORDS HTG; CpG island.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE Tracey, A.
AUTHORS Direct Submission
TITLE Submitted (20-OCT-2000) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
JOURNAL requests: clonerequest@sanger.ac.uk
COMMENT On Aug 29, 2000 this sequence version replaced gi:9843926. During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above. This sequence has been finished according to sequence map criteria as follows. An attempt is made to resolve all sequencing problems, such as compressions and repeats, but not necessarily within known annotated human repeat sequence elements (e.g. Alu). Where the sequence is ambiguous, there is an annotation using the 'unsure' feature key. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: En.: EMBL; Sw.: SWISSPROT; Tr.: TrEMBL; Wp.: WormPEP; Information on the WormPEP database can be found at http://www.sanger.ac.uk/Projects/C-elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping Group. Further information can be found at http://www.sanger.ac.uk/HGP/Chr6 RP11-239H6 is from the library RPC1-11.1 constructed at the Roswell Park Cancer Institute by the group of Pieter de Jong. For further details see http://bacpac.med.buffalo.edu/VECTOR: pBACe3.6
FEATURES
    source
        1..173456
            /organism="Homo sapiens"
            /mol_type="genomic DNA"
            /db_xref="taxon:9606"
            /chromosome="6"
            /clone="RP11-239H6"
            /clone_lib="RPC1-11.1"
            1609..2046
            /note="LTR16B repeat: matches 1..464 of consensus"
            2688..2713
    repeat_region
    repeat_region
```

repeat_region /note="13 copies 2 mer tt 92% conserved"
4959. 5048
/note="L1M1 repeat: matches 5492. 5581 of consensus"
5508. 5533
/note="13 copies 2 mer ac 100% conserved"
complement(5984. 6363)
/note="match: GSS: Em:AQ05521"
8132. 8622
/note="MLT2CB repeat: matches 1. 501 of consensus"
11067. 11517
/note="match: GSS: Em:AQ086344"
11303. 11541
/note="L1P13 repeat: matches 5881. 6140 of consensus"
complement(11581. 12028)
/note="match: GSS: Em:AQ464397"
/note="match: GSS: Em:AQ486628"
12033. 12675
/note="match: GSS: Em:AQ239805"
12035. 12449
/note="match: GSS: Em:AQ0828417"
12037. 12660
/note="match: GSS: Em:AQ236775"
12041. 12600
/note="match: GSS: Em:AQ489650"
12056. 12447
/note="match: GSS: Em:AQ045010"
12299. 12387
/note="MER5B repeat: matches 83. 178 of consensus"
12875. 13212
/note="MLT1A1 repeat: matches 11. 365 of consensus"
14882. 15037
/note="THER1B repeat: matches 1. 364 of consensus"
15462. 15516
/note="L1M63 repeat: matches 6063. 6116 of consensus"
18055. 18255
/note="L1M61 repeat: matches 5930. 6121 of consensus"
complement(19123. 19461)
/note="match: GSS: Em:AQ564070"
23382. 23776
/note="MER93 repeat: matches 7. 368 of consensus"
24079. 24106
/note="MER93 repeat: matches 368. 397 of consensus"
24500. 24637
/note="L1P repeat: matches 4241. 4385 of consensus"
24630. 24754
/note="L1P repeat: matches 5109. 5233 of consensus"
25993. 26923
/note="L1M66 repeat: matches 4769. 5706 of consensus"
26933. 27145
/note="Charlie3 repeat: matches 1. 215 of consensus"
27146. 27585
/note="L1M66 repeat: matches 5719. 6300 of consensus"
28238. 28386
/note="L1 repeat: matches 5151. 5308 of consensus"
28765. 29297
/note="MER1A repeat: matches 1. 527 of consensus"
29622. 30025
/note="MLT1B repeat: matches 1. 390 of consensus"
complement(31073. 31648)
/note="match: GSS: Em:AQ345866"
31965. 32002
/note="19 copies 2 mer ac 97% conserved"
32293. 32487
/note="MER5B repeat: matches 7. 176 of consensus"
32550. 32908
/note="MER47A repeat: matches 1. 366 of consensus"
35097. 35318
/note="L1M6 repeat: matches 5371. 5603 of consensus"
36107. 36248
/note="AluSg/x repeat: matches 133. 274 of consensus"
36249. 36614
/note="183 copies 2 mer ga 66% conserved"

repeat_region 37398. 37431
/note="17 copies 2 mer aa 85% conserved"
40629. 41044
/note="MSTA repeat: matches 1. 426 of consensus"
42019. 42574
/note="MER7C repeat: matches 1. 602 of consensus"
46407. 47179
/note="match: GSS: Em:AQ385283"
46739. 46926
/note="L1P10 repeat: matches 5958. 6153 of consensus"
47112. 47740
/note="L1P10 repeat: matches 5365. 5987 of consensus"
49919. 50232
/note="MLT1A1 repeat: matches 1. 374 of consensus"
51550. 51903
/note="THER1A repeat: matches 1. 354 of consensus"
complement(52137. 52547)
/note="match: GSS: Em:AQ016629"
52671. 52702
/note="16 copies 2 mer gt 100% conserved"
53069. 53456
/note="MLT2FB repeat: matches 1. 412 of consensus"
complement(56179. 56798)
/note="match: GSS: Em:AQ267603"
56189. 56750
/note="MER61E repeat: matches 1. 569 of consensus"
56796. 57235
/note="match: GSS: Em:AQ137177"
57071. 57187
/note="L1M8 repeat: matches 6123. 6238 of consensus"
57354. 57759
/note="L1M8 repeat: matches 5713. 6123 of consensus"
58839. 59181
/note="MSD repeat: matches 1. 394 of consensus"
61949. 62748
/note="L1M3 repeat: matches 6961. 7734 of consensus"
63038. 63383
/note="L1M3 repeat: matches 6632. 6961 of consensus"
complement(63091. 63520)
/note="match: GSS: Em:AQ818541"
63188. 63781
/note="match: GSS: Em:AQ485404"
63396. 63431
/note="18 copies 2 mer tg 94% conserved"
65033. 65652
/note="MER4A repeat: matches 2. 660 of consensus"
66656. 67052
/note="MER57B repeat: matches 1. 403 of consensus"
68264. 68774
/note="match: GSS: Em:AQ885163"
69704. 69789
/note="43 copies 2 mer aa 68% conserved"
70719. 71158
/note="match: GSS: Em:AQ805723"
71459. 71576
/note="MER57-internal repeat: matches 7151. 7270 of consensus"
75291. 75453
/note="MER58 repeat: matches 28. 196 of consensus"
77757. 77898
/note="LTR33 repeat: matches 364. 513 of consensus"
79760. 79799
/note="20 copies 2 mer ac 95% conserved"
79828. 80692
/note="CpG island"
/evidence=not experimental
81464. 81505
/note="21 copies 2 mer ac 76% conserved"

Query Match 3.6%; Score 204.4; DB 9; Length 173456;
Best Local Similarity 72.3%; Pred. No. 5.2e-29;
Matches 279; Conservative 0; Mismatches 106; Indels 1; Gaps 1;

```

Qy 3264 TTTAAATTTTTTAAAAAATAGATGAGTGGTGGCTCATGCCCTGTAATCCCAACACTTT 3323
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
19809 TCTAATTTTAAAGAAAAAATGGGCTGGGATGGTGGCTCATGCCCTGTAATCCTAACACTTT 19868
Qy 3324 GGGAGCCGGTCCGGAGGATAGCTTTCAGTCCAGCAGTTTGAGACCAGTCAGGGCAACAC 3383
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
19869 GGGAGGCCAAGGGCGATGGATCGCTGAGCTCAGGCGTTTGCACCCAGCCTGGGCAACAC 19928
Qy 3384 AGCAAGACCCCATATCTTAAAAAACAACAAAAAATACTCTGGGTATGGTTGTGCT 3443
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
19929 GGTGAACCCCGTCTCTACTAAAAAATAACAAAAAATTTACCTGGCATGGTGTG 19988
Qy 3444 CACCTGTAGTCCAGCTACACAGAAAGCTGAGGAGAGGATCACTTGAGCCCGAGGAGT 3503
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
19989 CACCTGTAATCCAGCGACTCAAGAGGCTGAGGAGGAGATCACTTGAACCCAGGAGGC 20048
Qy 3504 TGAGGCTGCAGTGATCCATGAACGCGCTGTACTACT-CAGTCTGGGTGACAGTGCAGAA 3562
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
20049 GGAGGTCGCAGTGAGCCAAAGATCCGCCACTGACACTACAGCCTGGGCGACAGAGCGAGAC 20108
Qy 3563 GCTGTCTCAAAAATAATAATAATAATAATACTTTTAAAAAACAATAATAATAATAAAT 3622
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
20109 TCCGTCTCAAAAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAAT 20168
Qy 3623 TTTAAAAACAACACACTAGAGATG 3648
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
20169 TAATTAATTTAAAAAAGAAAAAGTG 20194

```

Search completed: June 23, 2005, 07:51:11
Job time : 23476 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model
Run on: June 22, 2005, 14:13:22 ; Search time 2718 Seconds
(without alignments)
12362.197 Million cell updates/sec

Title: US-10-030-294-1
Perfect score: 5676
Sequence: 1 cagctgttcaggatgctg.....tttataataattgttcctcg 5676

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues
Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16Dec04:*
1: Geneseq1980s:*
2: Geneseq1990s:*
3: Geneseq2000s:*
4: Geneseq2001as:*
5: Geneseq2001bs:*
6: Geneseq2002as:*
7: Geneseq2002bs:*
8: Geneseq2003as:*
9: Geneseq2003bs:*
10: Geneseq2003cs:*
11: Geneseq2003ds:*
12: Geneseq2004as:*
13: Geneseq2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5676	100.0	5676	5	AAF25904 Human tum
C 2	5676	100.0	5676	5	AAF25907 Human tum
C 3	5676	100.0	5676	5	AAF25906 Human tum
4	5676	100.0	5960	5	AAF25905 Human tum
5	5673.4	100.0	13940	5	AAF25915 Human tum
C 6	196.2	3.5	58326	12	ADO48537 Human man
C 7	196	3.5	304905	11	ADP75180 Human aut
C 8	193.4	3.4	118384	10	ABX56555 Human aut
9	189.4	3.3	1384	3	AC74312 Human sec
10	188.4	3.3	135005	12	ADQ19501 Human sof
11	187.6	3.3	8133	6	ABN99663 Human clu
12	186	3.3	18861	12	ADJ12659 DNA fragm
13	185.4	3.3	47999	8	AD52898 Human tve
C 14	184.8	3.3	199878	10	ADL13719 Osteoarth
C 15	183	3.2	335199	10	ADC24703 Human wil
16	182.6	3.2	47903	13	ABD33591 Human can
17	182.2	3.2	36651	6	AAD28072 Human kin
C 18	181.8	3.2	49339	9	ADB16928 Human DYX
19	181	3.2	58985	9	ABZ59738 Human sec
20	181	3.2	143601	8	ABZ22654 Human epi

C	21	180.8	3.2	198522	11	ACN44010	Acn44010	Human gen
	22	180	3.2	26329	4	AAS42041	Aas42041	Genomic s
	23	180	3.2	94719	10	ADE95902	Ades95902	Human STA
	24	180	3.2	94720	9	ADA02654	Ada02654	Human STA
	25	180	3.2	94720	10	ADB72392	Adb72392	Human STA
	26	179.8	3.2	108316	10	ADC87336	Adc87336	Human GPC
C	27	179.6	3.2	65608	6	ABL62910	AbL62910	Breast ca
	28	179.6	3.2	65608	6	ABL64414	AbL64414	Stomach c
	29	179.6	3.2	65608	6	ABL67668	AbL67668	Oesophagu
C	30	179.4	3.2	214520	10	ADL13471	Adl13471	Osteoarth
C	31	179.2	3.2	94330	11	ACN44662	Acn44662	Human gen
C	32	178.6	3.1	186957	8	ACN56071	Acn56071	Human sec
C	33	178.6	3.1	201986	11	ACN44430	Acn44430	Human gen
	34	178.4	3.1	47219	13	ADR16283	Adr16283	Human Cay
	35	178.4	3.1	59884	13	ADR16284	Adr16284	Human Cay
C	36	178.2	3.1	630	4	AAL15941	Aal15941	Human bre
C	37	178.2	3.1	906	11	ACN85902	Acn85902	Breast ca
	38	178.2	3.1	7726	5	ABA21079	AbA21079	Human ner
	39	178.2	3.1	25423	4	AAK90279	Aak90279	Human dig
	40	178.2	3.1	25423	4	AAI57656	Aai57656	Human col
	41	178.2	3.1	25423	6	ABS99833	Abs99833	Genomic D
	42	178.2	3.1	25423	10	ADB92986	Adb92986	Human col
	43	178.2	3.1	25424	4	AAK90280	Aak90280	Human dig
	44	178.2	3.1	25424	4	AAI57657	Aai57657	Human col
	45	178.2	3.1	25424	6	ABS99834	Abs99834	Genomic D

ALIGNMENTS

RESULT 1
AAF25904
ID AAF25904 standard; DNA; 5676 BP.
XX
AC AAF25904;
XX
DT 19-APR-2001 (first entry)
XX
DE Human tumor suppressor gene p51 promoter associated DNA SEQ ID 1.
XX
KW Tumor suppressor; p51; cell death; cell proliferation; cancer;
KW cytosstatic; gene therapy; screening; db.
XX
OS Homo sapiens.
XX
PN W0200100818-A1.
XX
PD 04-JAN-2001.
XX
PF 28-JUN-2000; 2000WO-JP004261.
XX
PR 29-JUN-1999; 99JP-00181195.
XX
PA (NIPK) NIPPON KAYAKU KK.
XX
PI (SAKA/) SAKAI T.
XX
PI Sakai T, Kagaya S, Sato T, Sukenaga Y, Fujii H;
XX
DR WPI; 2001-112452/12.
XX
PT New DNA sequences encoding the human p51 promoter domain for screening
PT for compounds that modify p51 promoter activity and for diagnosing and
PT treating cancer.
XX
PS Claim 1(1); Page 31-34; 60pp; Japanese.
XX
CC This invention describes novel DNA sequences (I) encoding the human p51
CC promoter domain (which may include the 5'-untranslated sequence) or
CC derived from it by addition, deletion and/or substitution of one or more
CC bases. The invention also describes (1) expression plasmids including (I)
CC ; (2) host cells transformed by (1); (3) DNA probes binding to all or
CC part of (1); (4) cloning (I) using (3); (5) DNA (and RNA corresponding to
CC it) inhibiting the expression of p51, which includes the antisense

CC sequence of (1); (6) screening compounds for their ability to modify p51
CC promoter activity, by observing their effect on cells transformed by (1);
CC (7) compounds identified by (6); and (8) drug compositions containing
CC (7). The products of the invention have cytostatic activity and can be
CC used for gene therapy. (1) is used to screen for compounds that modify
CC p51 promoter activity. (1) and RNA corresponding to it are used to
CC inhibit the expression of p51. (1) and the identified compounds are used
CC for the diagnosis and treatment of cancer
XX
SQ Sequence 5676 BP; 1784 A; 1053 C; 1116 G; 1723 T; 0 U; 0 Other;

Query Match 100.0%; Score 5676; DB 5; Length 5676;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 5676; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	CAGCTGTTGAGGATGTCGGAAGAAGAGCCCAACCAATGCTCTCGACACCTGGGTGT	60
DB	1	CAGCTGTTGAGGATGTCGGAAGAAGAGCCCAACCAATGCTCTCGACACCTGGGTGT	60
QY	61	GACTTTGGAGGTATCAGGTTGCTGTTAAAGAACTGCCAACCCTCTCCGCCCAAT	120
DB	61	GACTTTGGAGGTATCAGGTTGCTGTTAAAGAACTGCCAACCCTCTCCGCCCAAT	120
QY	121	TGGCCTCTGTTCCCTTGCACTGCCCTCTTCTTGGGACACTCCCTTAAAGGCATCTTCTTG	180
DB	121	TGGCCTCTGTTCCCTTGCACTGCCCTCTTCTTGGGACACTCCCTTAAAGGCATCTTCTTG	180
QY	181	ACATTAACTAACTAATAATGTTTATTTGATGAATTTCACTGAGCCTGAAGAGATGGAG	240
DB	181	ACATTAACTAACTAATAATGTTTATTTGATGAATTTCACTGAGCCTGAAGAGATGGAG	240
QY	241	GTCAATCAAGAAGACACATGCTAGGTTGCAATGCACTGCTTTTCAATGAATTA	300
DB	241	GTCAATCAAGAAGACACATGCTAGGTTGCAATGCACTGCTTTTCAATGAATTA	300
QY	301	AGTCATTGCAATACCAATTCAGTTTACTTAAAGTTCTAGGCCAGCTTTACTCTTAATCGAT	360
DB	301	AGTCATTGCAATACCAATTCAGTTTACTTAAAGTTCTAGGCCAGCTTTACTCTTAATCGAT	360
QY	361	GTGAGCTGTAGCAATATTTAGTTCGAAGTTGGAGAGTTAGAGATCCTCTCCATGA	420
DB	361	GTGAGCTGTAGCAATATTTAGTTCGAAGTTGGAGAGTTAGAGATCCTCTCCATGA	420
QY	421	CAGAACTTTGGCTTCCACTTTTACTTAAATAGAGATTGTGTGGTTGAGCTGCAGCTATGA	480
DB	421	CAGAACTTTGGCTTCCACTTTTACTTAAATAGAGATTGTGTGGTTGAGCTGCAGCTATGA	480
QY	481	CAGAAAGTGTCTATACAAATTTAAATACCAAACTCAGTCTCTTCAATTTGAGCAATAGT	540
DB	481	CAGAAAGTGTCTATACAAATTTAAATACCAAACTCAGTCTCTTCAATTTGAGCAATAGT	540
QY	541	TGTTGAATTTACTCCACCACCTCTCTCTTGAAGTTCTTCTGCTCTCTCACTATA	600
DB	541	TGTTGAATTTACTCCACCACCTCTCTCTTGAAGTTCTTCTGCTCTCTCACTATA	600
QY	601	AATGCAGATGACTGGAAGGCTAGGACCTGAGTTTCAGTTACCTTGACACAAAGGAAT	660
DB	601	AATGCAGATGACTGGAAGGCTAGGACCTGAGTTTCAGTTACCTTGACACAAAGGAAT	660
QY	661	TCAGTTCTCTGATCTCATAGTCACAGGCTGCCAGAGCTCTACGGAACTGCAAGATCAT	720
DB	661	TCAGTTCTCTGATCTCATAGTCACAGGCTGCCAGAGCTCTACGGAACTGCAAGATCAT	720
QY	721	CTGCTTTAAGCCTCTGTGTGGGATCTGTGTTTCCAGTCCCTGTACCTATTGCTCT	780
DB	721	CTGCTTTAAGCCTCTGTGTGGGATCTGTGTTTCCAGTCCCTGTACCTATTGCTCT	780
QY	781	TTCCCTTTGTTAAAGAACCTTTATTTCTCTGAAACTCTCTGCTCAGTCATGGTAGGG	840
DB	781	TTCCCTTTGTTAAAGAACCTTTATTTCTCTGAAACTCTCTGCTCAGTCATGGTAGGG	840
QY	841	CCATCAGTCCACATGATCAGGCTCTCTGCGCAAAACATGGCATCTTTCTTTTGGGAAT	900
DB			

DB	841	CCATCAGTCCACATGATCAGGCTCTCTGCGCAAAACATGGCATCTTTCTTTTGGGAAT	900
QY	901	TGAATCTTAAAGCTGAATAGCTGAAGTTCAAAAAGCTGTTGAATCTGATTAAGCTTAC	960
DB	901	TGAATCTTAAAGCTGAATAGCTGAAGTTCAAAAAGCTGTTGAATCTGATTAAGCTTAC	960
QY	961	AGTGGCTTTGCAAAAGTGAAGTCTCAATCTTCTTAAAGTCCCTGAAATTAATTTAT	1020
DB	961	AGTGGCTTTGCAAAAGTGAAGTCTCAATCTTCTTAAAGTCCCTGAAATTAATTTAT	1020
QY	1021	CTGTGTACAGCCCTTTCTGAGATGTGTTTCTTTTCCAACTGCTCTTATAGTCTGT	1080
DB	1021	CTGTGTGTACAGCCCTTTCTGAGATGTGTTTCTTTTCCAACTGCTCTTATAGTCTGT	1080
QY	1081	GAAATTTTCAATATCTTTTTCATACATTTTTCATGTTTGTGTTGTGTTGTTT	1140
DB	1081	GAAATTTTCAATATCTTTTTCATACATTTTTCATGTTTGTGTTGTGTTGTTT	1140
QY	1141	GGCTTTAGGTAGGCAAGATCAGTTCTGTTGTTTATACCAAGGAATCTGTATGATACA	1200
DB	1141	GGCTTTAGGTAGGCAAGATCAGTTCTGTTGTTTATACCAAGGAATCTGTATGATACA	1200
QY	1201	TCCTTCCCTTTAAAAATAAAGTATCTAAAGCTCAAAAGAGTAGGCTACCTGCCTGAGG	1260
DB	1201	TCCTTCCCTTTAAAAATAAAGTATCTAAAGCTCAAAAGAGTAGGCTACCTGCCTGAGG	1260
QY	1261	TCGTGGAGTAAAGTAGTACAGAGCTCGTACTAAACCCAGGTTAGCCAACTGCTTTACAC	1320
DB	1261	TCGTGGAGTAAAGTAGTACAGAGCTCGTACTAAACCCAGGTTAGCCAACTGCTTTACAC	1320
QY	1321	AACATTTGCTCTCTCTTCAGAGTTATAGCAGTCTTTGGAAGAAAGAGTACTATTTTGC	1380
DB	1321	AACATTTGCTCTCTCTTCAGAGTTATAGCAGTCTTTGGAAGAAAGAGTACTATTTTGC	1380
QY	1381	CAAGA CCTCAGGAGGACCAAGAACAAAGTTCTGGGATATGTGATGATGAACCTTTAA	1440
DB	1381	CAAGAACCTCAGGAGGACCAAGAACAAAGTTCTGGGATATGTGATGATGAACCTTTAA	1440
QY	1441	AGTTTGTGGA CTCTGGCCATTAATGTGTATCTAAGACAGATTTTCTTATAGCT	1500
DB	1441	AGTTTGTGGA CTCTGGCCATTAATGTGTATCTAAGACAGATTTTCTTATAGCT	1500
QY	1501	AAACAAACAAACAAGAGATCCACAGGTTCCAGCAGCTATAATAAGAGTGAATTAAGTAC	1560
DB	1501	AAACAAACAAACAAGAGATCCACAGGTTCCAGCAGCTATAATAAGAGTGAATTAAGTAC	1560
QY	1561	AGTTGCAACATGAATATATCTCAGAAACCATGGCATCAATGAGCAAAAAAATCCAGAC	1620
DB	1561	AGTTGCAACATGAATATATCTCAGAAACCATGGCATCAATGAGCAAAAAAATCCAGAC	1620
QY	1621	ACAGAGATAGTACCATATGCTGCTGATTTATGTGATATCTAGCATTTGATTTGTTT	1680
DB	1621	ACAGAGATAGTACCATATGCTGCTGATTTATGTGATATCTAGCATTTGATTTGTTT	1680
QY	1681	CATAGTAGCAGCTAGCCTCAATGCTGCTTCAAAATTAAGTTGATTTAAATTAAGTAAGA	1740
DB	1681	CATAGTAGCAGCTAGCCTCAATGCTGCTTCAAAATTAAGTTGATTTAAATTAAGTAAGA	1740
QY	1741	ATAAAAAATTTAGCTCTTCTAGTAGCGTTAGCCCATGTAACTAGTGGCTACCAATCAGAC	1800
DB	1741	ATAAAAAATTTAGCTCTTCTAGTAGCGTTAGCCCATGTAACTAGTGGCTACCAATCAGAC	1800
QY	1801	GGTGCAAAATATAGATATTTTCTTTATACAGAAAGTTCTATTTGGAACAATGTTCTAG	1860
DB	1801	GGTGCAAAATATAGATATTTTCTTTATACAGAAAGTTCTATTTGGAACAATGTTCTAG	1860
QY	1861	AAAAATATACATATCTATAAAAAACAAAGCAAGTCACTGATTTCTTAAGGCCAGGGG	1920
DB	1861	AAAAATATACATATCTATAAAAAACAAAGCAAGTCACTGATTTCTTAAGGCCAGGGG	1920
QY	1921	TCAGGGGAGATCGATTTGCAAAAGTGTGTATGAGAAAGTTTGGGGTAAATAGGGTGTGGA	1980
DB	1921	TCAGGGGAGATCGATTTGCAAAAGTGTGTATGAGAAAGTTTGGGGTAAATAGGGTGTGGA	1980

Qy	1981	ATCTTGATTCGCGATGAAGGCTACTCGGTGTCTAATGTGTCACTCCTCAGACTGAACACT	2040
Db	1981	ATCTTGATTCGCGATGAAGGCTACTCGGTGTCTAATGTGTCACTCCTCAGACTGAACACT	2040
Qy	2041	TGGAATTTGGCGAATTTTCATTTGATGTAAATTTACCTCATAAAGTAACTCTAAGAGGTCA	2100
Db	2041	TGGAATTTGGCGAATTTTCATTTGATGTAAATTTACCTCATAAAGTAACTCTAAGAGGTCA	2100
Qy	2101	AGTGTGTTTGGGAAATTTATTTTAACTAGTTCGAATCTTATTTATGAGATGATTTTGGCA	2160
Db	2101	AGTGTGTTTGGGAAATTTATTTTAACTAGTTCGAATCTTATTTATGAGATGATTTTGGCA	2160
Qy	2161	AATACATAAAACATCTTATTCATTCATTTAGTTCGAATTTTGTGTAGTCTCCTGAAACAC	2220
Db	2161	AATACATAAAACATCTTATTCATTCATTTAGTTCGAATTTTGTGTAGTCTCCTGAAACAC	2220
Qy	2221	AGAGATGAATTTAGAAATAGCAAGCCTGCGCTCAAGCTGTTCAATCCAGTACAGGAGATG	2280
Db	2221	AGAGATGAATTTAGAAATAGCAAGCCTGCGCTCAAGCTGTTCAATCCAGTACAGGAGATG	2280
Qy	2281	AGTCTATTCAAAATAGCTAGACTCCAGGAGAAAGTTATAGTGACCTTACACAAAAA	2340
Db	2281	AGTCTATTCAAAATAGCTAGACTCCAGGAGAAAGTTATAGTGACCTTACACAAAAA	2340
Qy	2341	GTGCGAGATATATTTATGTAGGACAGTAAAGTGGGAGAGTTTCTTTTATGTGGAATAA	2400
Db	2341	GTGCGAGATATATTTATGTAGGACAGTAAAGTGGGAGAGTTTCTTTTATGTGGAATAA	2400
Qy	2401	GAGGAGAAATTTTGGTCTTTGAAGGATGAGCAAGTGTGAATGCGCAGATGGAGTTT	2460
Db	2401	GAGGAGAAATTTTGGTCTTTGAAGGATGAGCAAGTGTGAATGCGCAGATGGAGTTT	2460
Qy	2461	TAAACAAATTCCTGTGGAGGCGAGAATATGATCCAGGCAACAGAGCAACACAGAAAAA	2520
Db	2461	TAAACAAATTCCTGTGGAGGCGAGAATATGATCCAGGCAACAGAGCAACACAGAAAAA	2520
Qy	2521	TGCAACTAGAGAAAGTGCATGAGGGGAGCAGTGTGAATATTTTCATGAATGTAA	2580
Db	2521	TGCAACTAGAGAAAGTGCATGAGGGGAGCAGTGTGAATATTTTCATGAATGTAA	2580
Qy	2581	GTGAGAGAAATTTCTATCATAGACACCTGAGTTTGGCAGAGTGCATGTTCTTGGCTCCTA	2640
Db	2581	GTGAGAGAAATTTCTATCATAGACACCTGAGTTTGGCAGAGTGCATGTTCTTGGCTCCTA	2640
Qy	2641	GGAGTCAAGAGAAACAAAGTGTCCCTTCTCCTAGTTATGCTCAGTGGTCCAAAGTCCAA	2700
Db	2641	GGAGTCAAGAGAAACAAAGTGTCCCTTCTCCTAGTTATGCTCAGTGGTCCAAAGTCCAA	2700
Qy	2701	AACACCTTTCCTTCTTAAGTACTTTCTTCTCCCTCCATACAAATCTTAAGTCTTCACA	2760
Db	2701	AACACCTTTCCTTCTTAAGTACTTTCTTCTCCCTCCATACAAATCTTAAGTCTTCACA	2760
Qy	2761	AACATCATTTAAACAGGCGAGTCAATGTCAGAAAGGCAATTTCTTCTCCTAGACTCTAT	2820
Db	2761	AACATCATTTAAACAGGCGAGTCAATGTCAGAAAGGCAATTTCTTCTCCTAGACTCTAT	2820
Qy	2821	GTAGCTTATTTATATATATTTCTGCTTAAAGACTCTTAAAGTCTTGGAAAGTTTCCAC	2880
Db	2821	GTAGCTTATTTATATATATTTCTGCTTAAAGACTCTTAAAGTCTTGGAAAGTTTCCAC	2880
Qy	2881	CTTGACATCAAGATATATATTCATGCAATTTGTATAGTAACTTAGTCCCTTAAGAGAT	2940
Db	2881	CTTGACATCAAGATATATATTCATGCAATTTGTATAGTAACTTAGTCCCTTAAGAGAT	2940
Qy	2941	AAGGATGAACATATAAATAGAGTAAATTTGTTAAATATTAATATGATTCGCCACTTAT	3000
Db	2941	AAGGATGAACATATAAATAGAGTAAATTTGTTAAATATTAATATGATTCGCCACTTAT	3000
Qy	3001	TTTCACTTGTGCTGTATGTTGATGCTACTGTGTTGTTCTGTGAAATTTCTAGAGAGTTTG	3060
Db	3001	TTTCACTTGTGCTGTATGTTGATGCTACTGTGTTGTTCTGTGAAATTTCTAGAGAGTTTG	3060

Qy	3061	CCTCTTTTTCTGGGTCAACTCTGGCCATTTATTTCCATAATGCAATAGAGCCAATCTT	3120
Db	3061	CCTCTTTTTCTGGGTCAACTCTGGCCATTTATTTCCATAATGCAATAGAGCCAATCTT	3120
Qy	3121	TTTCATAATTACTTATTTAAATTTGTCGCCATTTAAATTTCTGTTCTCTTAGCTTAGTA	3180
Db	3121	TTTCATAATTACTTATTTAAATTTGTCGCCATTTAAATTTCTGTTCTCTTAGCTTAGTA	3180
Qy	3181	ACTTTAGGATTTTAAATAACCAACTATTGAAATCATGACATACGTTTAAATGATATTTAT	3240
Db	3181	ACTTTAGGATTTTAAATAACCAACTATTGAAATCATGACATACGTTTAAATGATATTTAT	3240
Qy	3241	TAAATACGTTAGGCTATAAAACCTTTTAAATTTTAAAAAATAGATGAGTGTGGCT	3300
Db	3241	TAAATACGTTAGGCTATAAAACCTTTTAAATTTTAAAAAATAGATGAGTGTGGCT	3300
Qy	3301	CATGCTGTAAATCCCAACACTTTTGGGAGCCGGTCCGGAGGATAGCTTGAGTCCAGCAG	3360
Db	3301	CATGCTGTAAATCCCAACACTTTTGGGAGCCGGTCCGGAGGATAGCTTGAGTCCAGCAG	3360
Qy	3361	TTTGAGACAGTCCAGGCAACACAGCAAGACCCCATATCTTAAAAAACAACAAACAA	3420
Db	3361	TTTGAGACAGTCCAGGCAACACAGCAAGACCCCATATCTTAAAAAACAACAAACAA	3420
Qy	3421	AATTTACCTGGTATGTTGTGCTCACCTGTAGTCCAGGTACACAGGAAGCTGAGGCGA	3480
Db	3421	AATTTACCTGGTATGTTGTGCTCACCTGTAGTCCAGGTACACAGGAAGCTGAGGCGA	3480
Qy	3481	AGGATCATTTGAGCCAGAGGTTGAGGCTGAGTGCATGATCCATGAACGGCTGTACACTC	3540
Db	3481	AGGATCATTTGAGCCAGAGGTTGAGGCTGAGTGCATGATCCATGAACGGCTGTACACTC	3540
Qy	3541	AGTCTGGGTGACAGTGCAGAACTGTCTCAAAAAATAAATAAATAAATAAATAAATAA	3600
Db	3541	AGTCTGGGTGACAGTGCAGAACTGTCTCAAAAAATAAATAAATAAATAAATAAATAA	3600
Qy	3601	AAAAAACAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA	3660
Db	3601	AAAAAACAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA	3660
Qy	3661	TTATTTGGGAGTCTATATCCCTGGAAGTTAAATTTAAATAATTTAGAGAGTCTTCTCA	3720
Db	3661	TTATTTGGGAGTCTATATCCCTGGAAGTTAAATTTAAATAATTTAGAGAGTCTTCTCA	3720
Qy	3721	TTTCTTAGACCTCGAATTTGTAATATCAGAGCTAGAAGCAACACTAGGCTGCCACT	3780
Db	3721	TTTCTTAGACCTCGAATTTGTAATATCAGAGCTAGAAGCAACACTAGGCTGCCACT	3780
Qy	3781	CCAAAGTGTGCTCCAGGACCCAGCAGCATCAAGTAACTCTGGGAACTGTTAGAAATGCG	3840
Db	3781	CCAAAGTGTGCTCCAGGACCCAGCAGCATCAAGTAACTCTGGGAACTGTTAGAAATGCG	3840
Qy	3841	AGTCTTAGGCTCACTCCAGACCTTCTGAAACCAAGATCTGCAATTAACAAGATTTCTAGT	3900
Db	3841	AGTCTTAGGCTCACTCCAGACCTTCTGAAACCAAGATCTGCAATTAACAAGATTTCTAGT	3900
Qy	3901	GCCTCAGGCAACATTAATAAATTTGAGAGCTCTGCACTAGAAATCTTCACTCCACCTTC	3960
Db	3901	GCCTCAGGCAACATTAATAAATTTGAGAGCTCTGCACTAGAAATCTTCACTCCACCTTC	3960
Qy	3961	ATTATAAATGGAATCACTTGGCTGTGGTCAACAGAAATTTGATTTTAAATTTAGAA	4020
Db	3961	ATTATAAATGGAATCACTTGGCTGTGGTCAACAGAAATTTGATTTTAAATTTAGAA	4020
Qy	4021	CTTCTTATTTAGGTCATCTATATTTGCTAATAGAGGAAAGCAAACTCTTTAACT	4080
Db	4021	CTTCTTATTTAGGTCATCTATATTTGCTAATAGAGGAAAGCAAACTCTTTAACT	4080
Qy	4081	GCAATTAACAAATCTATAATTAATAGTAAAGCAATCTTCCCTTAAAGTTTACATTTG	4140
Db	4081	GCAATTAACAAATCTATAATTAATAGTAAAGCAATCTTCCCTTAAAGTTTACATTTG	4140
Qy	4141	TGGAGCAAGCTGTTGATTTGGCTGGGCTCAGGCCGCTGTTGTTGTAATTTTCAAT	4200

Db 4141 TGGAGCAAGCTGTTTGAATTTGGCTGGGGCTCAGCGCGCCTGTTTGTGAATTTCACAATT 4200
 Qy 4201 CACAGATGTTAGCGGCTCTCGGGTAAGTAAGAAGAGATGTCAGATTTTAAATAGCT 4260
 Db 4201 CACAGATGTTAGCGGCTCTCGGGTAAGTAAGAAGAGATGTCAGATTTTAAATAGCT 4260
 Qy 4261 TCTCCCTTCATCTCGGCTGAAGCAACAATAATTTTATGAAACACATTTTGAGT 4320
 Db 4261 TCTCCCTTCATCTCGGCTGAAGCAACAATAATTTTATGAAACACATTTTGAGT 4320
 Qy 4321 TAGATTTACTACAGGGAATGTCAAAATTTCTCTGAAAGGGCTTTAGATTTCTCAAC 4380
 Db 4321 TAGATTTACTACAGGGAATGTCAAAATTTCTCTGAAAGGGCTTTAGATTTCTCAAC 4380
 Qy 4381 TTTGACATCTACTGATGTCACCTATTTACAGGTGTGTCCTGTGACTAGGGGGTGAAGGA 4440
 Db 4381 TTTGACATCTACTGATGTCACCTATTTACAGGTGTGTCCTGTGACTAGGGGGTGAAGGA 4440
 Qy 4441 AGATGTGAACCTCACATGTTAGTGACCGTTAGATACACAGAGTGGTTTTTTTCCCCCTG 4500
 Db 4441 AGATGTGAACCTCACATGTTAGTGACCGTTAGATACACAGAGTGGTTTTTTTCCCCCTG 4500
 Qy 4501 TTGGAGTCTATCTTAACCTGAGCTTCTGAATCATATTTTCAATTTCCAAATCCACAA 4560
 Db 4501 TTGGAGTCTATCTTAACCTGAGCTTCTGAATCATATTTTCAATTTCCAAATCCACAA 4560
 Qy 4561 ACCAGGATAGTTTACAGCCCATATTAGAAAGGAATAATTTATTTTGTGTGAGACTT 4620
 Db 4561 ACCAGGATAGTTTACAGCCCATATTAGAAAGGAATAATTTATTTTGTGTGAGACTT 4620
 Qy 4621 TCCTGATATTACATCTGATTTGGGAATATATGAACAATTTTATGTTTCCCTTCGAGTAG 4680
 Db 4621 TCCTGATATTACATCTGATTTGGGAATATATGAACAATTTTATGTTTCCCTTCGAGTAG 4680
 Qy 4681 GTCAAGTCAAAGCAAAACCAAAACAGCAAAACCTGTAAGACATTAAGATAGAGTGGAG 4740
 Db 4681 GTCAAGTCAAAGCAAAACCAAAACAGCAAAACCTGTAAGACATTAAGATAGAGTGGAG 4740
 Qy 4741 CCGACTCAGAGATTAATAATACTAGAAATTTTATTAACAGGCAATTTGAAATTAATTT 4800
 Db 4741 CCGACTCAGAGATTAATAATACTAGAAATTTTATTAACAGGCAATTTGAAATTAATTT 4800
 Qy 4801 GTGCACCTCAGAAATTTCTCAATAATATATTTTCCAAATTTTAAATCTTTAAAGAAA 4860
 Db 4801 GTGCACCTCAGAAATTTCTCAATAATATATTTTCCAAATTTTAAATCTTTAAAGAAA 4860
 Qy 4861 TTACTATATTATATGTAAGTACATGTGCATGTGTTTGAGGTAGGATATTTAACTCAATA 4920
 Db 4861 TTACTATATTATATGTAAGTACATGTGCATGTGTTTGAGGTAGGATATTTAACTCAATA 4920
 Qy 4921 AGTTATTTTCTTTTATTCGGGTCAGCAAGCTTCTAAGGGGATGTGAAGGGATATCT 4980
 Db 4921 AGTTATTTTCTTTTATTCGGGTCAGCAAGCTTCTAAGGGGATGTGAAGGGATATCT 4980
 Qy 4981 CTTTCTCTAGCTCAGAGGAAGTGTAGTCTCTAGTTTAAATATAATCAAGGAATTTCCCT 5040
 Db 4981 CTTTCTCTAGCTCAGAGGAAGTGTAGTCTCTAGTTTAAATATAATCAAGGAATTTCCCT 5040
 Qy 5041 GTCTTTCTATTGAGATTTGTACCAACAACAGCGGTTGGCTGAAAGGGAAAACCTGAAGGG 5100
 Db 5041 GTCTTTCTATTGAGATTTGTACCAACAACAGCGGTTGGCTGAAAGGGAAAACCTGAAGGG 5100
 Qy 5101 CGGGGAGGGGGAATAGATGAAAAAACAACAACAAACCTCCCTTAAGAGCTCT 5160
 Db 5101 CGGGGAGGGGGAATAGATGAAAAAACAACAACAAACCTCCCTTAAGAGCTCT 5160
 Qy 5161 ACAAAACATTTAGCCCCAGAAATAGTCAGAAATCCTCAATCAACAGATATCCAGA 5220
 Db 5161 ACAAAACATTTAGCCCCAGAAATAGTCAGAAATCCTCAATCAACAGATATCCAGA 5220
 Qy 5221 TACAAGGAAGTGTATAGCTGAGCAGGGTGGACACTCATCAGCTCAGTTTCAAGTTACA 5280

Db 5221 TACAAGGAAGTGTATAGCTGAGCAGGGTGGACACTCATCAGCTCAGTTTCAAGTTACA 5280
 Qy 5281 AAGTCCAGGCTGCTGAAATTAACCTCTGATGCCATTTATGCCAGCATCCAATCAGGACA 5340
 Db 5281 AAGTCCAGGCTGCTGAAATTAACCTCTGATGCCATTTATGCCAGCATCCAATCAGGACA 5340
 Qy 5341 GAGATCAGAAGTTCAGAGATGCTCCAGCTCCAAATTTGCCAAACAACAAGTGTGGCTACTA 5400
 Db 5341 GAGATCAGAAGTTCAGAGATGCTCCAGCTCCAAATTTGCCAAACAACAAGTGTGGCTACTA 5400
 Qy 5401 TAGCTCAAGGACTCTGAAGCCGTGAGAGAGGGGGAAGAAACAACAGTAGAGAGATGCCCA 5460
 Db 5401 TAGCTCAAGGACTCTGAAGCCGTGAGAGAGGGGGAAGAAACAACAGTAGAGAGATGCCCA 5460
 Qy 5461 GCTGGTAAGAAATCGAGTGTATGAATTTTATGAATCAATCTCATTTGGCTTAAA 5520
 Db 5461 GCTGGTAAGAAATCGAGTGTATGAATTTTATGAATCAATCTCATTTGGCTTAAA 5520
 Qy 5521 TCAAGAAACGCTCCGCTCTTTTGCATAATATGTAAGAGAGAGAGTGCCTTAACTTCTA 5580
 Db 5521 TCAAGAAACGCTCCGCTCTTTTGCATAATATGTAAGAGAGAGAGTGCCTTAACTTCTA 5580
 Qy 5581 TGTCTGATAGCAATTTGACCCCTATGCTTTTAGCTCCCGCTTTATATCTATATATACAC 5640
 Db 5581 TGTCTGATAGCAATTTGACCCCTATGCTTTTAGCTCCCGCTTTATATCTATATATACAC 5640
 Qy 5641 AGGTATTTGTATATTTTATATATATTTTCTCCGT 5676
 Db 5641 AGGTATTTGTATATTTTATATATATTTTCTCCGT 5676

RESULT 2
 AAF25907/c
 ID AAF25907 standard; RNA; 5676 BP.
 XX
 AC AAF25907;
 XX
 DT 19-APR-2001 (first entry)
 XX
 DE Human tumor suppressor gene p51 promoter associated RNA SEQ ID 4.
 XX
 KW Tumor suppressor; p51; cell death; cell proliferation; cancer;
 KW cytostatic; gene therapy; screening; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO200100818-A1.
 XX
 PD 04-JAN-2001.
 XX
 PF 28-JUN-2000; 2000WO-JP004261.
 XX
 PR 29-JUN-1999; 99JP-0018195.
 XX
 PA (NIPK) NIPPON KAYAKU KK.
 PA (SAKA/) SAKAI T.
 XX
 PI Sakai T, Kagaya S, Sato T, Sukenaga Y, Fujii H;
 XX
 WIPI; 2001-112452/12.
 XX
 PT New DNA sequences encoding the human p51 promoter domain for screening
 PT for compounds that modify p51 promoter activity and for diagnosing and
 PT treating cancer.
 XX
 PS Disclosure; Page 42-45; 60pp; Japanese.
 XX
 CC This invention describes novel DNA sequences (I) encoding the human p51
 CC promoter domain (which may include the 5'-untranslated sequence) or
 CC derived from it by addition, deletion and/or substitution of one or more
 CC bases. The invention also describes (1) expression plasmids including (1)
 CC ; (2) host cells transformed by (1); (3) DNA probes binding to all or
 CC part of (I); (4) cloning (I) using (3); (5) DNA (and RNA corresponding to

Db 3756 TGAGGGGAGATCGAATTCGAAAGTGGTATGAGGAAAGTTTTGGGGTAAATAGGGTTGTTGGA 3697
Qy 1981 ATCTTTGATTCGATGAAGGCTACTCGGTGCTCTAATGTGTCACCTCCCTCAGACGTGAACACT 2040
Db 3696 ATCTTGATTTGGGATGAAGGCTACTCGGTGCTCTAATGTGTCACCTCCCTCAGACGTGAACACT 3637
Qy 2041 TGGAAATTTGGGAAATTTCAATTTGATGTAAATTTATACCTCATAAAGTAACTCTAAGAGGTCA 2100
Db 3636 TGGAAATTTGGGAAATTTCAATTTGATGTAAATTTATACCTCATAAAGTAACTCTAAGAGGTCA 3577
Qy 2101 AGTGTTTGGGAAATTTATTTTAAATCAGTTGCAATACCTTATATAGAGATGATTTTGGCA 2160
Db 3576 AGTGTTTGGGAAATTTATTTTAAATCAGTTGCAATACCTTATATAGAGATGATTTTGGCA 3517
Qy 2161 AATACATAACATGTTATTCATTCATTTAGGTGCAATATTTTGTGCTAGCTCCTGAAACAC 2220
Db 3516 AATACATAACATGTTATTCATTCATTTAGGTGCAATATTTTGTGCTAGCTCCTGAAACAC 3457
Qy 2221 AGAGATGAAATTTAGAAATAGCAAGCTCGCCCTCAAGCTGTTCCACAATCCAGTACAGGAGATG 2280
Db 3456 AGAGATGAAATTTAGAAATAGCAAGCTCGCCCTCAAGCTGTTCCACAATCCAGTACAGGAGATG 3397
Qy 2281 AGTCTATTTCAAAAATAGCTAGATCTCCAGGAGAAAGTTATAGGTGACCTTTACACAAAAA 2340
Db 3396 AGTCTATTTCAAAAATAGCTAGATCTCCAGGAGAAAGTTATAGGTGACCTTTACACAAAAA 3337
Qy 2341 GTGCAGATATAATTTATGTAGGACATGAGAAAGTGTGAAGTGTGAAGTGTGAGTGTGCGAAGAAA 2400
Db 3336 GTGCAGATATAATTTATGTAGGACATGAGAAAGTGTGAAGTGTGAGTGTGCGAAGAAA 3277
Qy 2401 GAGGAGAAATTTTGGTCTTTTGAAGGATGAGCAAGATGTGAATGTGCGAGATGGAGTTT 2460
Db 3276 GAGGAGAAATTTTGGTCTTTTGAAGGATGAGCAAGATGTGAATGTGCGAGATGGAGTTT 3217
Qy 2461 TAAACATTCCTGTTGAGGCGAGAAATATCATCCAGGCAAGAGCAACCCAGAAATA 2520
Db 3216 TAAACATTCCTGTTGAGGCGAGAAATATCATCCAGGCAAGAGCAACCCAGAAATA 3157
Qy 2521 TGCACCTTAGGNAAGTGCATGAGGGGAGCAGTGTGTAAATAATTTTCATGAATGTAA 2580
Db 3156 TGCACCTTAGGNAAGTGCATGAGGGGAGCAGTGTGTAAATAATTTTCATGAATGTAA 3097
Qy 2581 GTGAGAGAAATTTGTATCATAGACACCTGAGTTTGGCAGAGTGCATGTTCTTGCTCCTA 2640
Db 3096 GTGAGAGAAATTTGTATCATAGACACCTGAGTTTGGCAGAGTGCATGTTCTTGCTCCTA 3037
Qy 2641 GGAGTCAAGAGAAACAAAGTGTCCCTTTCCTACAGTTATGCTCAGTGTGCCAAGTCCAA 2700
Db 3036 GGAGTCAAGAGAAACAAAGTGTCCCTTTCCTACAGTTATGCTCAGTGTGCCAAGTCCAA 2977
Qy 2701 AACACCTTTCCTTCAAGTACTTTCTTCTCCCTCCATACAAATCTAAGTCTTCACA 2760
Db 2976 AACACCTTTCCTTCAAGTACTTTCTTCTCCCTCCATACAAATCTAAGTCTTCACA 2917
Qy 2761 AACATCATTTTAAACAGGCAGTATGCTCAGAAAGCAATGTGCTTTCCTCCTCCTAT 2820
Db 2916 AACATCATTTTAAACAGGCAGTATGCTCAGAAAGCAATGTGCTTTCCTCCTCCTAT 2857
Qy 2821 GTAGTTTATTTATATACAAATTTCTGCTTAAAGTCTTAAAGTCTTGGAAAAGTTTCCAC 2880
Db 2856 GTAGTTTATTTATATACAAATTTCTGCTTAAAGTCTTAAAGTCTTGGAAAAGTTTCCAC 2797
Qy 2881 CTTGCAATCAAGATATAATTCATGCAATTTGTATAGTAACTTAGTCCCTTAAGAGAT 2940
Db 2796 CTTGCAATCAAGATATAATTCATGCAATTTGTATAGTAACTTAGTCCCTTAAGAGAT 2737
Qy 2941 AAGGATCAACTATAATAATAAGAGTAAATTTATGTAATTAATAATGATTTGCCACTTAT 3000
Db 2736 AAGGATCAACTATAATAATAAGAGTAAATTTATGTAATTAATAATGATTTGCCACTTAT 2677
Qy 3001 TTTTCACTTGTATGTTGATGCTACTCGGTGTTCTGTTGAAATTTCTAGAGAGTTTG 3060
Db 2676 TTTTCACTTGTATGTTGATGCTACTCGGTGTTCTGTTGAAATTTCTAGAGAGTTTG 2617

Qy 3061 CCTCTTTTCTGGGTCAAATCTCGCAATTTATTTCCATAATGCAATAGGAGCAATCTT 3120
Db 2616 CCTCTTTTCTGGGTCAAATCTCGCAATTTATTTCCATAATGCAATAGGAGCAATCTT 2557
Qy 3121 TTTTCAFAATTTACTTATTTTAAAAATTTTGTGCAATTTTAAATTTCTCTCTTAGCTTAGTA 3180
Db 2556 TTTTCAFAATTTACTTATTTTAAAAATTTTGTGCAATTTTAAATTTCTCTCTTAGCTTAGTA 2497
Qy 3181 ACTTTAGGATTTTAAATTAACAACTATTGAAATCATGACATACGTTTAAATGATATTT 3240
Db 2496 ACTTTAGGATTTTAAATTAACAACTATTGAAATCATGACATACGTTTAAATGATATTT 2437
Qy 3241 TAAATACGTTTAGCTATAAACCTTTTAAATTTTAAAAAATAGATGAGTGTGGTGGCT 3300
Db 2436 TAAATACGTTTAGCTATAAACCTTTTAAATTTTAAAAAATAGATGAGTGTGGTGGCT 2377
Qy 3301 CATGCCGTAAATCCCAACACTTTTGGGAAGCCGGTGGGAGGATAGCTTTGAGTCCAGCAG 3360
Db 2376 CATGCCGTAAATCCCAACACTTTTGGGAAGCCGGTGGGAGGATAGCTTTGAGTCCAGCAG 2317
Qy 3361 TTTGAGACCACTCAGGGCAACACAGCAAGACCCCATATCTTAAAAAACHAAAAAACHAA 3420
Db 2316 TTTGAGACCACTCAGGGCAACACAGCAAGACCCCATATCTTAAAAAACHAAAAAACHAA 2257
Qy 3421 AATTACCTCGGTATGTTGTTGCTCACTGTGTAGTCCAAAGCTACACAGGAAGCTGAGGCAGA 3480
Db 2256 AATTACCTCGGTATGTTGTTGCTCACTGTGTAGTCCAAAGCTACACAGGAAGCTGAGGCAGA 2197
Qy 3481 AGGATCACTTGAGCCCAAGAGGTTGAGGCTGAGTGCATCCATGAAACGCGCTGCTACACTC 3540
Db 2196 AGGATCACTTGAGCCCAAGAGGTTGAGGCTGAGTGCATCCATGAAACGCGCTGCTACACTC 2137
Qy 3541 AGTCTGGGTGACAGTGCAAGAGCTGCTCTCAAAAAATAATAATAATAATAATACTTTT 3600
Db 2136 AGTCTGGGTGACAGTGCAAGAGCTGCTCTCAAAAAATAATAATAATAATAATACTTTT 2077
Qy 3601 AAAAAAACHAAAAATTAATAATAATAATAATAATAATAATAATAATAATAATAATACTTCA 3660
Db 2076 AAAAAAACHAAAAATTAATAATAATAATAATAATAATAATAATAATAATAATAATACTTCA 2017
Qy 3661 TTAATTTGGGAGTCTATATATCCCTGGAGATTTAAATTTAAAAATAATTTAGAAAGTCTTCTCTCA 3720
Db 2016 TTAATTTGGGAGTCTATATATCCCTGGAGATTTAAATTTAAAAATAATTTAGAAAGTCTTCTCTCA 1957
Qy 3721 TTTTCTAGAGAGCTCGAATTTGAAATATCAGAGCTAGAAAGGAACACTAGGGCTCGCCACT 3780
Db 1956 TTTTCTAGAGAGCTCGAATTTGAAATATCAGAGCTAGAAAGGAACACTAGGGCTCGCCACT 1897
Qy 3781 CCAAGTGTGGTCCAGGACCAAGCAGCATCAAGTAACTGGGAAACGTTTAGAAATGCAAG 3840
Db 1896 CCAAGTGTGGTCCAGGACCAAGCAGCATCAAGTAACTGGGAAACGTTTAGAAATGCAAG 1837
Qy 3841 AGTCTTAGGCTCACCCAGACCTACTGAAACCAAGATCTGCAATTAACAAGATTTCTTAGGT 3900
Db 1836 AGTCTTAGGCTCACCCAGACCTACTGAAACCAAGATCTGCAATTAACAAGATTTCTTAGGT 1777
Qy 3901 GCTCAAGGCAACATTAATACTTTGAGAGCTCTGCACTAGAAATCTTCACTCCACCTTTC 3960
Db 1776 GCTCAAGGCAACATTAATACTTTGAGAGCTCTGCACTAGAAATCTTCACTCCACCTTTC 1717
Qy 3961 ATTATAAATGGAATCACTTGGGCTGTGGTCAAGGAAATTTGATTTTAAATTTTCAGAA 4020
Db 1716 ATTATAAATGGAATCACTTGGGCTGTGGTCAAGGAAATTTGATTTTAAATTTTCAGAA 1657
Qy 4021 CCTTCTATTTAGTCACTATATTTTGTAAATAGCAGGAGGAAAGCAAACTCTTTAACT 4080
Db 1656 CCTTCTATTTAGTCACTATATTTTGTAAATAGCAGGAGGAAAGCAAACTCTTTAACT 1597
Qy 4081 GCAATTAACAAATCTATAATTAATTTAGTTAAGCAATCTTCCCTTTAAGTTTTCATCTTTG 4140
Db 1596 GCAATTAACAAATCTATAATTAATTTAGTTAAGCAATCTTCCCTTTAAGTTTTCATCTTTG 1537

CC pert of (1); (4) cloning (1) using (3); (5) DNA (and RNA corresponding to
CC it) inhibiting the expression of p51, which includes the antisense
CC sequence of (1); (6) screening compounds for their ability to modify p51
CC promoter activity, by observing their effect on cells transformed by (1);
CC (7) compounds identified by (6); and (8) drug compositions containing
CC (7). The products of the invention have cytostatic activity and can be
CC used for gene therapy. (1) is used to screen for compounds that modify
CC p51 promoter activity. (1) and RNA corresponding to it are used to
CC inhibit the expression of p51. (1) and the identified compounds are used
CC for the diagnosis and treatment of cancer

XX SQ Sequence 5676 BP; 1723 A; 1116 C; 1053 G; 1784 T; 0 U; 0 Other;

Query Match 100.0%; Score 5676; DB 5; Length 5676;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 5676; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAGCTGTTGAGGATGTCGTGAAAGAGCCACCCACCATTCCTGAGACACTGGGTGT 60
DB 5676 CAGCTGTTGAGGATGTCGTGAAAGAGCCACCCACCATTCCTGAGACACTGGGTGT 5617
QY 61 GACTTTGGAGGATACAGGTTGCTGTTAAAGAACTGCAACCTCTTCTGCCCCAAT 120
DB 5616 GACTTTGGAGGATACAGGTTGCTGTTAAAGAACTGCAACCTCTTCTGCCCCAAT 5557
QY 121 TGGCTCTGTTCCCTTCGATGCCCTCTTCTTGGGACACTCCCTTAAGGCATCTTCTTG 180
DB 5556 TGGCTCTGTTCCCTTCGATGCCCTCTTCTTGGGACACTCCCTTAAGGCATCTTCTTG 5497
QY 181 ACATTAACCTTAACATATAATTTTATTTGATGAATTTCACTGACCTGGAAGAGATGGAG 240
DB 5496 ACATTAACCTTAACATATAATTTTATTTGATGAATTTCACTGACCTGGAAGAGATGGAG 5437
QY 241 GTCAATCAGAAAGACATGCGCTAAGGTTGCAATGCCATCTGCTTTTCAATGAATTA 300
DB 5436 GTCAATCAGAAAGACATGCGCTAAGGTTGCAATGCCATCTGCTTTTCAATGAATTA 5377
QY 301 AGTCATTCGAATACATTCAGTTTACATTAAGTTTCTAGGCCAGCTTACCTTAATCGAT 360
DB 5376 AGTCATTCGAATACATTCAGTTTACATTAAGTTTCTAGGCCAGCTTACCTTAATCGAT 5317
QY 361 GTCAGACTGTAGCAAAATATTAAGTCCAAAGTTGGAAGAGTTAGCAGGATCTCTCCATGA 420
DB 5316 GTCAGACTGTAGCAAAATATTAAGTCCAAAGTTGGAAGAGTTAGCAGGATCTCTCCATGA 5257
QY 421 CAGAACTTTGGCTTCCACTTTACTTAAATAGAGATTGTGTGGTTGAGCTGCAGCTATGTA 480
DB 5256 CAGAACTTTGGCTTCCACTTTACTTAAATAGAGATTGTGTGGTTGAGCTGCAGCTATGTA 5197
QY 481 CAGAAAGTGTACATATTAATAATCACCAACTCAGTCTCTTCAATTTGAGCAATAGT 540
DB 5196 CAGAAAGTGTACATATTAATAATCACCAACTCAGTCTCTTCAATTTGAGCAATAGT 5137
QY 541 TGGTGAATTTACTCCACACCTCTCTCTTCTTGAAGTTTCTTCTGCTCTCTCCCTCATATA 600
DB 5136 TGGTGAATTTACTCCACACCTCTCTCTTCTTGAAGTTTCTTCTGCTCTCTCCCTCATATA 5077
QY 601 AATGCAGATGACTGTGAAAGGCTAGGACCTGAGGTTCACTTCCCTGACACAAAGGAAT 660
DB 5076 AATGCAGATGACTGTGAAAGGCTAGGACCTGAGGTTCACTTCCCTGACACAAAGGAAT 5017
QY 661 TCAGTTTCTGATCTCATAGTCACAGGCTGCCAGAGCTCTACGGACATGCCAGATCAT 720
DB 5016 TCAGTTTCTGATCTCATAGTCACAGGCTGCCAGAGCTCTACGGACATGCCAGATCAT 4957
QY 721 CTGCTTTAAGCCTCTGTTGGTGGCATCTGTTGTTTTCCACTGCGCCTGTACTATTGCTCT 780
DB 4956 CTGCTTTAAGCCTCTGTTGGTGGCATCTGTTGTTTTCCACTGCGCCTGTACTATTGCTCT 4897
QY 781 TTCCTTGGTTAAAGAACACCTTTATTTTCTTGAAATCTCTGCTCAGTCATGGTAGGG 840
DB 4896 TTCCTTGGTTAAAGAACACCTTTATTTTCTTGAAATCTCTGCTCAGTCATGGTAGGG 4837

QY 841 CCATCAGTCCACATGATCAGGCTCTCTCTGGCCAAACATGCGATCTTCTTTTGGGAAT 900
DB 4836 CCATCAGTCCACATGATCAGGCTCTCTCTGGCCAAACATGCGATCTTCTTTTGGGAAT 4777
QY 901 TGAATCTTAAGCTGAATAGCTGAAGTTCAAAAAAGCTGTTGAATCTGACTTHAGCCTAC 960
DB 4776 TGAATCTTAAGCTGAATAGCTGAAGTTCAAAAAAGCTGTTGAATCTGACTTHAGCCTAC 4717
QY 961 AGTGGCTTTGCAAGGAGCTGCCATTCCTATTTCTTAAGTCCCTCAATTTATAATTTAT 1020
DB 4716 AGTGGCTTTGCAAGGAGCTGCCATTCCTATTTCTTAAGTCCCTCAATTTATAATTTAT 4657
QY 1021 CTTGGTTACAGCCCTTTCTCAGATGTGTGTTTTTTTTTCCAACTGCTCTCTATAGTCTGT 1080
DB 4656 CTTGGTTACAGCCCTTTCTCAGATGTGTGTTTTTTTTTCCAACTGCTCTCTATAGTCTGT 4597
QY 1081 GAATTTTCATATTTCTTTTCATACATTTTCATGTTTTGTTTTGTTTTGTTGTTGTTTTT 1140
DB 4596 GAATTTTCATATTTCTTTTCATACATTTTCATGTTTTGTTTTGTTTTGTTGTTGTTTTT 4537
QY 1141 GCGTTTTAGGTAGGCAATCAGTTTCTGTTTATACCCAGGAATCCTGATTTGATACA 1200
DB 4536 GCGTTTTAGGTAGGCAATCAGTTTCTGTTTATACCCAGGAATCCTGATTTGATACA 4477
QY 1201 TCTTCCCTTTAAAAATAAAGTATCTAAGGCTCAAGAGAGTAGGCTACCTGCCTGAGG 1260
DB 4476 TCTTCCCTTTAAAAATAAAGTATCTAAGGCTCAAGAGAGTAGGCTACCTGCCTGAGG 4417
QY 1261 TCTGGAGTAAGTTAGTACAGAGCTCGTACTAAACCCAGGTTAGCCAACTGCTTTACAC 1320
DB 4416 TCTGGAGTAAGTTAGTACAGAGCTCGTACTAAACCCAGGTTAGCCAACTGCTTTACAC 4357
QY 1321 AACATTTGCTCTCTCTTCCAGAGTTATAGCAGTCTTGGAGAAAGAGCTACTATTTTGC 1380
DB 4356 AACATTTGCTCTCTCTTCCAGAGTTATAGCAGTCTTGGAGAAAGAGCTACTATTTTGC 4297
QY 1381 CAAAGACCTCAGGAGACCAAGAAACAAAGTTCTGGGATATGTGATGATGCAACTCTTAAAA 1440
DB 4296 CAAAGACCTCAGGAGACCAAGAAACAAAGTTCTGGGATATGTGATGATGCAACTCTTAAAA 4237
QY 1441 AGTTTCTTGGACTTCTGGCCATAATTTGTATCTAAGACCAAGTTTCATTTCTTAATAGCT 1500
DB 4236 AGTTTCTTGGACTTCTGGCCATAATTTGTATCTAAGACCAAGTTTCATTTCTTAATAGCT 4177
QY 1501 AACAACAAACNAGAGATCCACAGGTTGAGCAGCTATTAATAGAGTGAATTTACTGATAC 1560
DB 4176 AACAACAAACNAGAGATCCACAGGTTGAGCAGCTATTAATAGAGTGAATTTACTGATAC 4117
QY 1561 AGTTGACAAATGAATATATCTCAGAAACCATGGCATCAATGAGCAAAAAAATCCAGAC 1620
DB 4116 AGTTGACAAATGAATATATCTCAGAAACCATGGCATCAATGAGCAAAAAAATCCAGAC 4057
QY 1621 ACAGAAAGATAGTACCATATGCTGCTCATTTATGTGATATTTCTAGCATTTGATTTGCCAA 1680
DB 4056 ACAGAAAGATAGTACCATATGCTGCTCATTTATGTGATATTTCTAGCATTTGATTTGCCAA 3997
QY 1681 CATAGTAGCCAGTAGCCTCACATGGCTATTCAAAATTAAGTTGATTTAAATTAAGTAAGA 1740
DB 3996 CATAGTAGCCAGTAGCCTCACATGGCTATTTCAAAATTAAGTTGATTTAAATTAAGTAAGA 3937
QY 1741 ATAAAAATTTAGCTCTTCTAGTAGCGTTAGCCACATGTAACTAGTGGCTACCATCAGAC 1800
DB 3936 ATAAAAATTTAGCTCTTCTAGTAGCGTTAGCCACATGTAACTAGTGGCTACCATCAGAC 3877
QY 1801 GGTGCAAAATAGAAATATTTTCTTTTATAACAGAAAGTCTTCTTTGGAAAAACAATGTTCTAG 1860
DB 3876 GGTGCAAAATAGAAATATTTTCTTTTATAACAGAAAGTCTTCTTTGGAAAAACAATGTTCTAG 3817
QY 1861 ABAATATACATATCTATATAAAAAAAGCAAGCTCAGTGAATTTCTTAAGCCAGGGG 1920
DB 3816 ABAATATACATATCTATATAAAAAAAGCAAGCTCAGTGAATTTCTTAAGCCAGGGG 3757
QY 1921 TGAGGGGAGATCGATTGTCACAAAGTGGTATGAGGAAAGTTTTTGGGGTAAATAGGGTTGTTGGA 1980

Qy	4141	TGAGCAAGCTGTTTGATTTTGGCTTGGGGCTCAGCCGGSCCTGTTTGTGAAATTTTCACAATT	4200
Db	1536	TGGAGCAAGCTGTTTGATTTTGGCTTGGGGCTCAGCCGGSCCTGTTTGTGAAATTTTCACAATT	1477
Qy	4201	CACAGATGTTAGCCGCTCTCTCGGGCTAAGTAAAGGAAGAGAAATGTCAAAGTTTAAATAGCT	4260
Db	1476	CACAGATGTTAGCCGCTCTCTCGGGCTAAGTAAAGGAAGAGAAATGTCAAAGTTTAAATAGCT	1417
Qy	4261	TCTCCCTTCCATCCTCGCTGAAGCAACAATAAATAATTTTATGAACAACATTTTGAGT	4320
Db	1416	TCTCCCTTCCATCCTCGCTGAAGCAACAATAAATAATTTTATGAACAACATTTTGAGT	1357
Qy	4321	TAGATTTACTTTACAGGGAATGTCAAATTTCTCTGAAAGGCTTTAGATGTCTCAAC	4380
Db	1356	TAGATTTACTTTACAGGGAATGTCAAATTTCTCTGAAAGGCTTTAGATGTCTCAAC	1297
Qy	4381	TTTGACATCTACTGATGTCACTATTTTACAGGTGTCTGTGTGATCAGGGGTGAAGGA	4440
Db	1296	TTTGACATCTACTGATGTCACTATTTTACAGGTGTCTGTGTGATCAGGGGTGAAGGA	1237
Qy	4441	AGATGTGAACCTCACCATGTTAGTCACCGTTAGATACACAGAGTGGTTTTTTTTCCCCCTG	4500
Db	1236	AGATGTGAACCTCACCATGTTAGTCACCGTTAGATACACAGAGTGGTTTTTTTTCCCCCTG	1177
Qy	4501	TTGGAGTCTATCTTAACTGAGCTTCTGAAATCATATTTTCATTCATTTCCAAATCCACAA	4560
Db	1176	TTGGAGTCTATCTTAACTGAGCTTCTGAAATCATATTTTCATTCATTTCCAAATCCACAA	1117
Qy	4561	ACCAGGATAAGTTTACAGCCCATTTTACAGGAAGAAATAATTTTGTGTGTAGACTT	4620
Db	1116	ACCAGGATAAGTTTACAGCCCATTTTACAGGAAGAAATAATTTTGTGTGTAGACTT	1057
Qy	4621	TCCTGATTTACACTGATTTTGGGAATATATGAACAAATTTTATGGTTTCCTTTCGAAGTAG	4680
Db	1056	TCCTGATTTACACTGATTTTGGGAATATATGAACAAATTTTATGGTTTCCTTTCGAAGTAG	997
Qy	4681	GTCAAGTCAAAGCAAAACCAAAAACAGCAAAAACGTGAAGACATAAAGATAGAGTGGAG	4740
Db	996	GTCAAGTCAAAGCAAAACCAAAAACAGCAAAAACGTGAAGACATAAAGATAGAGTGGAG	937
Qy	4741	CCGACTGAGAGATTTAAATAAATAACTAGAAATATTTTTATTAAACAGCAATTTTGAATAATTT	4800
Db	936	CCGACTGAGAGATTTAAATAAATAACTAGAAATATTTTTATTAAACAGCAATTTTGAATAATTT	877
Qy	4801	GTGCATTTCAGAATATCTCAATAATAATATTTATTCCAAATTTTAAATATCTTTAAGAAAA	4860
Db	876	GTGCATTTCAGAATATCTCAATAATAATATTTATTCCAAATTTTAAATATCTTTAAGAAAA	817
Qy	4861	TTACTATATATATGTAAGTACATGTGCATGTGTTTGGGTAGGATATTTAACTCAATAA	4920
Db	816	TTACTATATATATGTAAGTACATGTGCATGTGTTTGGGTAGGATATTTAACTCAATAA	757
Qy	4921	AGGTTATTTCTTTTATTTCCGGTCAAGCAAGCTTCTTAAGGGGATGTGAAGGGATATCT	4980
Db	756	AGGTTATTTCTTTTATTTCCGGTCAAGCAAGCTTCTTAAGGGGATGTGAAGGGATATCT	697
Qy	4981	CTTTCTCTTAGCTGAGAGGAAAGAGTGAGTCTCTAAGTTTAAATATATTAATCAAGGAATTTCCCT	5040
Db	696	CTTTCTCTTAGCTGAGAGGAAAGAGTGAGTCTCTAAGTTTAAATATATTAATCAAGGAATTTCCCT	637
Qy	5041	GTCTTTGCTATTTGAGATTGTGACCAACAGGGCGGTTGGCTGAAGGGAAACTGAAGGG	5100
Db	636	GTCTTTGCTATTTGAGATTGTGACCAACAGGGCGGTTGGCTGAAGGGAAACTGAAGGG	577
Qy	5101	CGGGAGGGAGGGAATAGATGAAAAACAAAAACAAAACTTCCCTAAGCAGCTCT	5160
Db	576	CGGGAGGGAGGGAATAGATGAAAAACAAAAACAAAACTTCCCTAAGCAGCTCT	517
Qy	5161	ACAAAAATTTTAGCCCCAGAAAAATAGTCAAGAAATCTCTCAATCAAAACGATATCAGA	5220
Db	516	ACAAAAATTTTAGCCCCAGAAAAATAGTCAAGAAATCTCTCAATCAAAACGATATCAGA	457

Qy	5221	TACAAGGAAGTGTATAGTACTGGAGCAGGGTGGACACTCATCAGCTCAGTTCAGTTTACA	5280
Db	456	TACAAGGAAGTGTATAGTACTGGAGCAGGGTGGACACTCATCAGCTCAGTTCAGTTTACA	397
Qy	5281	AAAGTCCAGGCTGCTGAAATTTAAACTCTGATGCCATTCATGCCAGCATCCAATCACGACA	5340
Db	396	AAAGTCCAGGCTGCTGAAATTTAAACTCTGATGCCATTCATGCCAGCATCCAATCACGACA	337
Qy	5341	GAGATCAGAAGTTTCAGAGATGCCTCCAGCTCCAAATTTGCCAAACAAGTGTGGCTACTA	5400
Db	336	GAGATCAGAAGTTTCAGAGATGCCTCCAGCTCCAAATTTGCCAAACAAGTGTGGCTACTA	277
Qy	5401	TAGCTCAAGGACTCTGAAAGCCGTCGAGAGAGGGGGAAGAACACAGTAGAGAGGATGCCCA	5460
Db	276	TAGCTCAAGGACTCTGAAAGCCGTCGAGAGAGGGGGAAGAACACAGTAGAGAGGATGCCCA	217
Qy	5461	GCTGGTAAAGAAATCGAGTGTATGAAGTTTTAGTCAATTGATGAATCTCATTGGCTAAAA	5520
Db	216	GCTGGTAAAGAAATCGAGTGTATGAAGTTTTAGTCAATTGATGAATCTCATTGGCTAAAA	157
Qy	5521	TCAAGAAACGCTCCGCCCTCTTTCGCAATATGATGAAGGAGAGAAGTGCCTAAACTTCTA	5580
Db	156	TCAAGAAACGCTCCGCCCTCTTTCGCAATATGATGAAGGAGAGAAGTGCCTAAACTTCTA	97
Qy	5581	TGCTCTGATAGCATTTGACCCCTATTGCTTTTAGCTCCCGGCTTTATATCTATATATACAC	5640
Db	96	TGCTCTGATAGCATTTGACCCCTATTGCTTTTAGCTCCCGGCTTTATATCTATATATACAC	37
Qy	5641	AGGTATTTGTGTATATTTTATATATAATTGTTCTCGGT	5676
Db	36	AGGTATTTGTGTATATTTTATATAAATTGTTCTCGGT	1
RESULT 4			
AAAF25905			
ID AAF25905 standard; DNA; 5960 BP.			
XX	AAAF25905;		
AC			
XX			
XX	19-APR-2001 (first entry)		
DT			
XX	Human tumor suppressor gene p51 promoter associated DNA SEQ ID 2.		
DE			
XX	Tumor suppressor; p51; cell death; cell proliferation; cancer;		
KW	cytostatic; gene therapy; screening; ds.		
KW			
XX	Homo sapiens.		
OS			
XX	WO200100818-A1.		
XX			
XX	04-JAN-2001.		
PD			
XX	28-JUN-2000; 2000WO-JP004261.		
PF			
XX	29-JUN-1999; 99JP-00183195.		
PR			
XX	(NIPK) NIPPON KAYAKU KK.		
PA			
PA	(SAKA/) SAKAI T.		
XX			
PI	Sakai T, Kagaya S, Sato T, Sukenaga Y, Fujii H;		
XX			
DR	WPI; 2001-112452/12.		
XX			
PT	New DNA sequences encoding the human p51 promoter domain for screening		
PT	for compounds that modify p51 promoter activity and for diagnosing and		
PT	treating cancer.		
XX			
PS	Claim 1(4); Page 35-38; 60pp; Japanese.		
XX			
XX	This invention describes novel DNA sequences (I) encoding the human p51		
CC	promoter domain (which may include the 5'-untranslated sequence) or		
CC	derived from it by addition, deletion and/or substitution of one or more		
CC	bases. The invention also describes (1) expression plasmids including (1)		

781 TTCCCTGGTTAACAGAACCTTTATTTCTTCTGAAAACCTCTCTGCTCAGTCATGGTAGGG 840

QY 1921 TGAGGGAGATCGATTGCAAGTGGTATGAGGAAAGTTTGGGTAATAGGGTTGTTGGA 1980
DB 1921 TGAGGGAGATCGATTGCAAGTGGTATGAGGAAAGTTTGGGTAATAGGGTTGTTGGA 1980
QY 1981 ATCTTGATTCGCATGAAAGGCTACTCGGTGTCTAATGTGTCCACCTCTCAGACTGAACACT 2040
DB 1981 ATCTTGATTCGCATGAAAGGCTACTCGGTGTCTAATGTGTCCACCTCTCAGACTGAACACT 2040
QY 2041 TGGNAITGGCGAATTCATTGTATGTAAATATACCTCATAAAGTAATCTTAAGAGTGCA 2100
DB 2041 TGGNAITGGCGAATTCATTGTATGTAAATATACCTCATAAAGTAATCTTAAGAGTGCA 2100
QY 2101 AGTGTITGGGAAATTAATTTTAACTCAGTTGGCAATACCTTATATGATGAGATGATTTTGCA 2160
DB 2101 AGTGTITGGGAAATTAATTTTAACTCAGTTGGCAATACCTTATATGATGAGATGATTTTGCA 2160
QY 2161 AATACATAAACATGTTATTTCATCCATTAGGTGCAATATTTTGTCTAGCTCCTGAAAACAC 2220
DB 2161 AATACATAAACATGTTATTTCATCCATTAGGTGCAATATTTTGTCTAGCTCCTGAAAACAC 2220
QY 2221 AGAGATGAATTAGAATAGCAAGCGCTGCCCTCAAGCTGTTCAACAATCAGTACAGGAGATG 2280
DB 2221 AGAGATGAATTAGAATAGCAAGCGCTGCCCTCAAGCTGTTCAACAATCAGTACAGGAGATG 2280
QY 2281 AGTCTATTCAAAAATAGCTAGACTCCAGGAGAAAGTTATAGGTGACTTACACAAAAA 2340
DB 2281 AGTCTATTCAAAAATAGCTAGACTCCAGGAGAAAGTTATAGGTGACTTACACAAAAA 2340
QY 2341 GTGCAGATATAATTTATGTAGGACAGTAGAAGTGGGAAAGTTCTTTTATGTGGAAAAA 2400
DB 2341 GTGCAGATATAATTTATGTAGGACAGTAGAAGTGGGAAAGTTCTTTTATGTGGAAAAA 2400
QY 2401 GAGGGAGAAATTTTGGTCTTTGAAGGATGAGCAAGATGTGAATATGCGCAGATGGAGTTT 2460
DB 2401 GAGGGAGAAATTTTGGTCTTTGAAGGATGAGCAAGATGTGAATATGCGCAGATGGAGTTT 2460
QY 2461 TAAACATTCCTGTGGAGGCGCAATATCATCCAGGCAAGAGCAACACAGAAAAA 2520
DB 2461 TAAACATTCCTGTGGAGGCGCAATATCATCCAGGCAAGAGCAACACAGAAAAA 2520
QY 2521 TGCNAACCTAGAGGAAAGTGCAAGGAGGAGCAGTTGTAAATAATTTTCATGAATGTAA 2580
DB 2521 TGCNAACCTAGAGGAAAGTGCAAGGAGGAGCAGTTGTAAATAATTTTCATGAATGTAA 2580
QY 2581 GTGAGAAGAAATTTGTATCATAGACACTGAGTTTGGCAGAGTGCAATGTTCTGGCTCCTA 2640
DB 2581 GTGAGAAGAAATTTGTATCATAGACACTGAGTTTGGCAGAGTGCAATGTTCTGGCTCCTA 2640
QY 2641 GGAGTCAAGAGAAACAAAGTGTCCCTTCTCCTACGTTATGCTCAGTGGTCCAAAGTCCAA 2700
DB 2641 GGAGTCAAGAGAAACAAAGTGTCCCTTCTCCTACGTTATGCTCAGTGGTCCAAAGTCCAA 2700
QY 2701 AACACCTTTCTCTTAAGTACTTTCTCTCCCTCCATACAAATCTAAAGTCTTCACA 2760
DB 2701 AACACCTTTCTCTTAAGTACTTTCTCTCCCTCCATACAAATCTAAAGTCTTCACA 2760
QY 2761 AACATCAATTAACAGCAGGTATGTCAGAAAGGCAATTTGCTTCTAGACTTCTAT 2820
DB 2761 AACATCAATTTAAACAGCAGGTATGTCAGAAAGGCAATTTGCTTCTAGACTTCTAT 2820
QY 2821 GTACGTTATTATATACAAATTTCTGCCCTAAAGACTCTAAAGTCTTGGAAAAAGTTTCCAC 2880
DB 2821 GTACGTTATTATATACAAATTTCTGCCCTAAAGACTCTAAAGTCTTGGAAAAAGTTTCCAC 2880
QY 2881 CTTCGACATCAAGATATAATTCATGCAATTTGTATAGTAACCTTAGTCCCTCAAGAGAAT 2940
DB 2881 CTTCGACATCAAGATATAATTCATGCAATTTGTATAGTAACCTTAGTCCCTCAAGAGAAT 2940
QY 2941 AAGGATGAACCTATAAATATAGAAAGTAATTTATGGTAAATTAATGATTTGCCACTTATT 3000
DB 2941 AAGGATGAACCTATAAATATAGAAAGTAATTTATGGTAAATTAATGATTTGCCACTTATT 3000
QY 3001 TTTTCACTTCGTGTGTTGATGCTACTGCTGTTCTGTTGAAATTTCTAGAGAGTTTG 3060

DB 3001 TTTTCACTTCGTGTGTTGATGCTACTGCTGTTCTGTTGAAATTTCTAGAGAGTTTG 3060
QY 3061 CTTCTTTTCTGGGTCAACTCTCGCAATTTATTTCCATAATGCAATAGAGGCAATCTT 3120
DB 3061 CTTCTTTTCTGGGTCAACTCTCGCAATTTATTTCCATAATGCAATAGAGGCAATCTT 3120
QY 3121 TTTTCATAATTACTTATTTAAATTTTGTGCAATTTATTTCTGTTCTCTTAGCTTAGTA 3180
DB 3121 TTTTCATAATTACTTATTTAAATTTTGTGCAATTTATTTCTGTTCTCTTAGCTTAGTA 3180
QY 3181 ACTTTAGGATTTTAAATAACCACTATTTGAAATCATGACATAGCTTTTAAATGATATATT 3240
DB 3181 ACTTTAGGATTTTAAATAACCACTATTTGAAATCATGACATAGCTTTTAAATGATATATT 3240
QY 3241 TAAATACGTTAGGCTATAAACCTTTTAAATTTTAAAAAATAGATGATGTTGGTGGCT 3300
DB 3241 TAAATACGTTAGGCTATAAACCTTTTAAATTTTAAAAAATAGATGATGTTGGTGGCT 3300
QY 3301 CATGCCCTGAATCCCAACACTTTTGGGAAGCCGGTCGGGAGGATAGCTTTGAGTCCAGCAG 3360
DB 3301 CATGCCCTGAATCCCAACACTTTTGGGAAGCCGGTCGGGAGGATAGCTTTGAGTCCAGCAG 3360
QY 3361 TTTGAGACCAAGTCAGGCGCAACACAGCAAGACCCCATATCTAAAAAACAACAAACAA 3420
DB 3361 TTTGAGACCAAGTCAGGCGCAACACAGCAAGACCCCATATCTAAAAAACAACAAACAA 3420
QY 3421 AATTACCTGGGTATGTTGTGCTCCTGTAGTCCAAAGCTACACAGGAAAGCTGAGGAG 3480
DB 3421 AATTACCTGGGTATGTTGTGCTCCTGTAGTCCAAAGCTACACAGGAAAGCTGAGGAG 3480
QY 3481 AGGATCACTTTGAGCCAGAGGTTGAGGCTGCAATGATCCATGAAACGCGCTGCTACACTC 3540
DB 3481 AGGATCACTTTGAGCCAGAGGTTGAGGCTGCAATGATCCATGAAACGCGCTGCTACACTC 3540
QY 3541 AGTCTGGGTGACAGTGCAAGAGCTGTCTCAAAAAATAAATAAATAAATAAATAAATA 3600
DB 3541 AGTCTGGGTGACAGTGCAAGAGCTGTCTCAAAAAATAAATAAATAAATAAATAAATA 3600
QY 3601 AAAAAACAAAAATTAATTTAAATTTTAAAAACAAACACTAGAGATGTTTGCAAATTTGA 3660
DB 3601 AAAAAACAAAAATTAATTTAAATTTTAAAAACAAACACTAGAGATGTTTGCAAATTTGA 3660
QY 3661 TTTATTTGGGAGTCTATATCCCTGGAAAGTTAAATTTTAAATAATTTAGAGAGTCTTCTCTCA 3720
DB 3661 TTTATTTGGGAGTCTATATCCCTGGAAAGTTAAATTTTAAATAATTTAGAGAGTCTTCTCTCA 3720
QY 3721 TTTTCTTAGAGAGCTCGAATTTGTAATATCAGAGCTAGAAAGGAACTAGGGCTCGCCACT 3780
DB 3721 TTTTCTTAGAGAGCTCGAATTTGTAATATCAGAGCTAGAAAGGAACTAGGGCTCGCCACT 3780
QY 3781 CCAAGTGTGGTCCAAAGGACCAAGCATCAAGTAACTCTGGGAAACGTTGTAGAAATGTCAG 3840
DB 3781 CCAAGTGTGGTCCAAAGGACCAAGCATCAAGTAACTCTGGGAAACGTTGTAGAAATGTCAG 3840
QY 3841 AGTCTTTAGGCTCACCCAGACCTTACTGAACCAAGATCTGCAATTAACAGATTTCTAGGT 3900
DB 3841 AGTCTTTAGGCTCACCCAGACCTTACTGAACCAAGATCTGCAATTAACAGATTTCTAGGT 3900
QY 3901 GCCTCAGCGGCACTTAAATCTTGAAGCTCTGCACTAGAAATCTTCACTCCACCTTTTC 3960
DB 3901 GCCTCAGCGGCACTTAAATCTTGAAGCTCTGCACTAGAAATCTTCACTCCACCTTTTC 3960
QY 3961 ATTTAAATTTGGAATCACTTTGGGCTGTGGTCACAGGAAATTTGATATTTTAAATTTTCA 4020
DB 3961 ATTTAAATTTGGAATCACTTTGGGCTGTGGTCACAGGAAATTTGATATTTTAAATTTTCA 4020
QY 4021 CTTTCTTATTAGTGTCTATATTTTCTTAATAGCAGGAGAAAGCAACTCTTTAACT 4080
DB 4021 CTTTCTTATTAGTGTCTATATTTTCTTAATAGCAGGAGAAAGCAACTCTTTAACT 4080
QY 4081 GCAATTAACAAATCTATAATTTAATAGTTAAGCAATCTTCCCTTTAAGTTTTCATATTTTG 4140

Db 4081 GCAATTAACAAATCTATAATTAATAGTTAAAGCAATCTTCCCTTTAAGTTTTACATTTTG 4140
Qy 4141 TGGAGCAAGCTGTTTGAATTTGGCTGGGCTCAGCGCGGCTGTTTGTGAATTTCACAATT 4200
Db 4141 TGGAGCAAGCTGTTTGAATTTGGCTGGGCTCAGCGCGGCTGTTTGTGAATTTCACAATT 4200
Qy 4201 CACAGATGTTAGCCGCTCTCGGCTAAGTAAAGGAAGAGAAATCTCAAGTTTAAATAGCT 4260
Db 4201 CACAGATGTTAGCCGCTCTCGGCTAAGTAAAGGAAGAGAAATCTCAAGTTTAAATAGCT 4260
Qy 4261 TCTCCCTTCCATCTCGCTGGAAGCAACAATAAATATTTTATGAACAACATTTTGAGT 4320
Db 4261 TCTCCCTTCCATCTCGCTGGAAGCAACAATAAATATTTTATGAACAACATTTTGAGT 4320
Qy 4321 TAGATTTACTTACAGGGAATGTCATTTCTCTGAAGGGCTTTAGATTTGTCTCAAC 4380
Db 4321 TAGATTTACTTACAGGGAATGTCATTTCTCTGAAGGGCTTTAGATTTGTCTCAAC 4380
Qy 4381 TTTGACATCTACTGATGTCACTTATTTACAGGTGTGTCTGTGACTAGGGGGTGAAGGGA 4440
Db 4381 TTTGACATCTACTGATGTCACTTATTTACAGGTGTGTCTGTGACTAGGGGGTGAAGGGA 4440
Qy 4441 AGATGTGAATCAACATGTTAGTGAACGTTAGATATACAGAGTGGTTTTTTTCCCGCTG 4500
Db 4441 AGATGTGAATCAACATGTTAGTGAACGTTAGATATACAGAGTGGTTTTTTTCCCGCTG 4500
Qy 4501 TTTGAGTCTATCTTAAGTCTGCTGATCTGATCATATTTCAATTTCCAAATCCCAAA 4560
Db 4501 TTTGAGTCTATCTTAAGTCTGCTGATCTGATCATATTTCAATTTCCAAATCCCAAA 4560
Qy 4561 ACCAGGATAAGTTTACAGCCCATATTTACAGAAAGGAATAAATTTATTTGTGTAGACTT 4620
Db 4561 ACCAGGATAAGTTTACAGCCCATATTTACAGAAAGGAATAAATTTATTTGTGTAGACTT 4620
Qy 4621 TCTCGATTTACTGATTTGGGAATATATGAACAATTTTATGGTTTCCCTTTGAGTAG 4680
Db 4621 TCTCGATTTACTGATTTGGGAATATATGAACAATTTTATGGTTTCCCTTTGAGTAG 4680
Qy 4681 GTCAAGTCAAGCAAAACCAAAAACAGCAAAACTGTAAAGACATAAAGATAGAGTGGAG 4740
Db 4681 GTCAAGTCAAGCAAAACCAAAAACAGCAAAACTGTAAAGACATAAAGATAGAGTGGAG 4740
Qy 4741 CCGACTGAGAGATTTAAATAAATAGATAATTTTATTAACAGCAATTTGAAATAATTT 4800
Db 4741 CCGACTGAGAGATTTAAATAAATAGATAATTTTATTAACAGCAATTTGAAATAATTT 4800
Qy 4801 GTGCACCTTCAGAAATTTCTCAATAATATATATTTTCCAAATTTTAATATCTTTAAGAAA 4860
Db 4801 GTGCACCTTCAGAAATTTCTCAATAATATATATTTTCCAAATTTTAATATCTTTAAGAAA 4860
Qy 4861 TTACTAT 4920
Db 4861 TTACTAT 4920
Qy 4921 AGTTATTTTCTTTTATTCGGCTCAGCAAGCTTCTAAGGGATGTGAAGGGATATCT 4980
Db 4921 AGTTATTTTCTTTTATTCGGCTCAGCAAGCTTCTAAGGGATGTGAAGGGATATCT 4980
Qy 4981 CTTTCTCTTAGCTGAGAGGAGAGTGTAGTCTTAAGTTTAAATATATATATATATATATATAT 5040
Db 4981 CTTTCTCTTAGCTGAGAGGAGAGTGTAGTCTTAAGTTTAAATATATATATATATATATATAT 5040
Qy 5041 GTCTTTCTATTTGAGATTTGTACCAACAGCGGTTGGCTGAAGGGAAACTGAAGGG 5100
Db 5041 GTCTTTCTATTTGAGATTTGTACCAACAGCGGTTGGCTGAAGGGAAACTGAAGGG 5100
Qy 5101 CGGGAGGGAGGGAATAGATGAATAAACAACAAACAAACAAACAAACAAACAAACAAACAAAC 5160
Db 5101 CGGGAGGGAGGGAATAGATGAATAAACAACAAACAAACAAACAAACAAACAAACAAACAAAC 5160
Qy 5161 ACAAACAATTTTAGCCCCAGAAATAGTACAGAAATCTCTCAAAATCAAAATCAAAATCAAAAT 5220
Db 5161 ACAAACAATTTTAGCCCCAGAAATAGTACAGAAATCTCTCAAAATCAAAATCAAAATCAAAAT

Qy 5221 TACAAGGAAGTCTTATGTAGCTGGAGCAGGGTGGACACTCATCAGCTCAGTTCAGTTACA 5280
Db 5221 TACAAGGAAGTCTTATGTAGCTGGAGCAGGGTGGACACTCATCAGCTCAGTTCAGTTACA 5280
Qy 5281 AAAGTCCAGGCTGCTGAATTTAAACTCTGATGCCATTCATGCCAGCATCCATCAGACA 5340
Db 5281 AAAGTCCAGGCTGCTGAATTTAAACTCTGATGCCATTCATGCCAGCATCCATCAGACA 5340
Qy 5341 GAGATCAGAAAGTTTCAGAGATGCTCCAGCTCCAAATTTGCCAAACAAGTGTGGCTACTA 5400
Db 5341 GAGATCAGAAAGTTTCAGAGATGCTCCAGCTCCAAATTTGCCAAACAAGTGTGGCTACTA 5400
Qy 5401 TACGTCAAGAGCTCTGAGAGCCGTGAGAGAGGGGGAAGAACACAGTAGAGAGATGCCCA 5460
Db 5401 TACGTCAAGAGCTCTGAGAGCCGTGAGAGAGGGGGAAGAACACAGTAGAGAGATGCCCA 5460
Qy 5461 GCTGGTAAGAATCGAGTGTATGAATTTAGTCAATTTAGTGAATTTCTATTGGCTAAAA 5520
Db 5461 GCTGGTAAGAATCGAGTGTATGAATTTAGTCAATTTAGTGAATTTCTATTGGCTAAAA 5520
Qy 5521 TCAAGAAACGCTCCGCTCTTTGCAAAATATGTATGAAGAGAGAGAGTGCCTAACTTCTA 5580
Db 5521 TCAAGAAACGCTCCGCTCTTTGCAAAATATGTATGAAGAGAGAGAGTGCCTAACTTCTA 5580
Qy 5581 TGTCTGATAGCATTTGACCCCTATTGCTTTTAGCTCCCGCTTTTATATCTATATATACAC 5640
Db 5581 TGTCTGATAGCATTTGACCCCTATTGCTTTTAGCTCCCGCTTTTATATCTATATATACAC 5640
Qy 5641 AGGTATTTGTATATTTTAT 5676
Db 5641 AGGTATTTGTATATTTTAT 5676

RESULT 5

AAF25915
ID AAF25915 standard; DNA; 13940 BP.

XX AAF25915;

XX 19-APR-2001 (first entry)

XX Human tumor suppressor gene p51 promoter DNA fragment SEQ ID 12.

XX Tumor suppressor; p51; cell death; cell proliferation; cancer;

XX Cytostatic; gene therapy; screening; ds.

XX Homo sapiens.

XX WO200100818-A1.

XX 04-JAN-2001.

XX 28-JUN-2000; 2000WO-JP004261.

XX 29-JUN-1999; 99JP-00183195.

XX (NIPK) NIPPON KAYAKU KK.

XX (SAKA/) SAKAI T.

XX Sakai T, Kagaya S, Sato T, Sukenaga Y, Fujii H;

XX WPI; 2001-112452/12.

XX New DNA sequences encoding the human p51 promoter domain for screening
PT for compounds that modify p51 promoter activity and for diagnosing and
PT treating cancer.

XX Example 1; Page 49-57; 60pp; Japanese.

XX This invention describes novel DNA sequences (I) encoding the human p51
CC promoter domain (which may include the 5'-untranslated sequence) or
CC derived from it by addition, deletion and/or substitution of one or more

CC bases. The invention also describes (1) expression plasmids including (1)
 CC ; (2) host cells transformed by (1); (3) DNA probes binding to all or
 CC part of (1); (4) cloning (1) using (3); (5) DNA (and RNA corresponding
 CC it) inhibiting the expression of p51, which includes the antisense
 CC sequence of (1); (6) screening compounds for their ability to modify p51
 CC promoter activity, by observing their effect on cells transformed by (1);
 CC (7) compounds identified by (6); and (8) drug compositions containing
 CC (7). The products of the invention have cytostatic activity and can be
 CC used for gene therapy. (1) is used to screen for compounds that modify
 CC p51 promoter activity. (1) and RNA corresponding to it are used to
 CC inhibit the expression of p51. (1) and the identified compounds are used
 CC for the diagnosis and treatment of cancer
 XX

SQ Sequence 13940 BP; 3949 A; 2972 C; 3085 G; 3934 T; 0 U; 0 Other;

Query Match 100.0%; Score 5673.4; DB 5; Length 13940;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 5674; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY	2	AGCTGTTTCAGGGATGCTGGAAGAAGCCACCCACATTCCTTGGACACTGGGTG	61
DB	51	ATCTGTTTCAGGGATGCTGGAAGAAGCCACCCACATTCCTTGGACACTGGGTG	110
QY	62	ACTTTGAGGGATATCAGGTTTGTCTGTTAAAGAACTGCCAACTCTTCTGCCCCAAT	121
DB	111	ACTTTGAGGGATATCAGGTTTGTCTGTTAAAGAACTGCCAACTCTTCTGCCCCAAT	170
QY	122	GGCTCTGTTCCCTTGGATCCCTCTTCTTGGGACACTCCCTTAAGGATCTTCTTGA	181
DB	171	GGCTCTGTTCCCTTGGATCCCTCTTCTTGGGACACTCCCTTAAGGATCTTCTTGA	230
QY	182	CATTAACTTAACATATAATGTTTATTTGATGAATTTTCAGTGACTGAAGAGAGATGGAG	241
DB	231	CATTAACTTAACATATAATGTTTATTTGATGAATTTTCAGTGACTGAAGAGAGATGGAG	290
QY	242	TCAATCAGAGAAGCAATGGCTTAAGTTGCAATGCACTTGCTTTTCAATTGAATTA	301
DB	291	TCAATCAGAGAAGCAATGGCTTAAGTTGCAATGCACTTGCTTTTCAATTGAATTA	350
QY	302	GTCAATTCGAATACCAATTCAGTTTACTTAAGTTCTTAGGCCAGCTTTACTCTTAATCGATG	361
DB	351	GTCAATTCGAATACCAATTCAGTTTACTTAAGTTCTTAGGCCAGCTTTACTCTTAATCGATG	410
QY	362	TCAGACTGTAGCAAAATTAAGTCCAAAGTTGGAAGGTTAGCAGGATCCTCTCCATGAC	421
DB	411	TCAGACTGTAGCAAAATTAAGTCCAAAGTTGGAAGGTTAGCAGGATCCTCTCCATGAC	470
QY	422	AGAACTTTGGCTTCCACTTTTACTAAATAGAGATTTGTGTGTTGAGCTGCAGCTATGTAC	481
DB	471	AGAACTTTGGCTTCCACTTTTACTAAATAGAGATTTGTGTGTTGAGCTGCAGCTATGTAC	530
QY	482	AGAAAGTGTATACAAATTAATAATACCAAACTCAGTCTCTTCAATTTGAGCAATAGTT	541
DB	531	AGAAAGTGTATACAAATTAATAATACCAAACTCAGTCTCTTCAATTTGAGCAATAGTT	590
QY	542	GGTGAATTTACTCCACCACTCTCTCTTGAAGTTCTTTCTGCTCTCTCACTATA	601
DB	591	GGTGAATTTACTCCACCACTCTCTCTTGAAGTTCTTTCTGCTCTCTCACTATA	650
QY	602	ATGCAGGATGACCTGGAAGGCTAGGACCTGAGTTTCAGTTTACCCTGACACAAAGGAAT	661
DB	651	ATGCAGGATGACCTGGAAGGCTAGGACCTGAGTTTCAGTTTACCCTGACACAAAGGAAT	710
QY	662	CAGTTTCTCTGATCTCATAGTACAGGCTGCGAGAGCTCTACGGAACATGCAAGATCATC	721
DB	711	CAGTTTCTCTGATCTCATAGTACAGGCTGCGAGAGCTCTACGGAACATGCAAGATCATC	770
QY	722	TGCTTTAAGCTCTTGTGGGATCTGTTGTTTCCACGCTGCTGACCTATGCTCTT	781
DB	771	TGCTTTAAGCTCTTGTGGGATCTGTTGTTTCCACGCTGCTGACCTATGCTCTT	830
QY	782	TCCTTGGTTAACAGAACTTTATTTCTTCTGAAAACTCTCTGCTCAGTCAATGGTAGGC	841

DB	831	TCCTTGGTTAACAGAACTTTATTTTCTTCTGAAAACTCTCTGCTCAGTCAATGGTAGGC	890
QY	842	CATCAGTCCACATGATCAGGCTCTCCTGCGCAAAACATGGCATCTTCTTTGGGAATTT	901
DB	891	CATCAGTCCACATGATCAGGCTCTCCTGCGCAAAACATGGCATCTTCTTTGGGAATTT	950
QY	902	GAATCTTAAGCTGAATAGCTGAAGTTCAAAAAAGCTGTTGAATCTGACTTACGCCTACA	961
DB	951	GAATCTTAAGCTGAATAGCTGAAGTTCAAAAAAGCTGTTGAATCTGACTTACGCCTACA	1010
QY	962	GTGGCTTTGCAAGAGTCTGCTCAATTCCTATTTCTTAAAGTCCCTGAAATTAATTAATC	1021
DB	1011	GTGGCTTTGCAAGAGTCTGCTCAATTCCTATTTCTTAAAGTCCCTGAAATTAATTAATC	1070
QY	1022	CTGGTTACAGCCCTTCTGAGATGCTGTTTCTTCCAACTGCTCTTATAGTCTGTG	1081
DB	1071	CTGGTTACAGCCCTTCTGAGATGCTGTTTCTTCCAACTGCTCTTATAGTCTGTG	1130
QY	1082	AATTTTCATATTTCTTTTCATACATTTTCATGTTTGTGTTTGTGTTGTTGTTGTTT	1141
DB	1131	AATTTTCATATTTCTTTTCATACATTTTCATGTTTGTGTTTGTGTTGTTGTTT	1190
QY	1142	GTCTTATAGTATGAGCAATTCAGTTTCTGTTGTTTATACCCAAAGAACTCTGATTTGAT	1201
DB	1191	GTCTTATAGTATGAGCAATTCAGTTTCTGTTGTTTATACCCAAAGAACTCTGATTTGAT	1250
QY	1202	CCTTCCCTTTTAAATAAAGTATCTAAGCTCAAGAGAGTAGGCTACTGCTGAGGT	1261
DB	1251	CCTTCCCTTTTAAATAAAGTATCTAAGCTCAAGAGAGTAGGCTACTGCTGAGGT	1310
QY	1262	CTGGAGTAGTAAGTTAGTACCAAGAGCTGTAATAACCCAGGTTAGCCAACTGCTTTTACACA	1321
DB	1311	CTGGAGTAGTAAGTTAGTACCAAGAGCTGTAATAACCCAGGTTAGCCAACTGCTTTTACACA	1370
QY	1322	ACATTTGCTCTCTCTTCTCAGAGTTATAGCAGTCTTGGAAAGAAAGAGCTACTATTTTGC	1381
DB	1371	ACATTTGCTCTCTCTTCTCAGAGTTATAGCAGTCTTGGAAAGAAAGAGCTACTATTTTGC	1430
QY	1382	AAAGACTCAGAGGACCAAGAACAGTTCTGGGATATCTGATGATTGAACTCTTAAAAA	1441
DB	1431	AAAGACTCAGAGGACCAAGAACAGTTCTGGGATATCTGATGATTGAACTCTTAAAAA	1490
QY	1442	GTTTGTGACACTCTTGGCCATAATTTGTGTATCTAAGCCAGATTTTCATTTTAATAGCTA	1501
DB	1491	GTTTGTGACACTCTTGGCCATAATTTGTGTATCTAAGCCAGATTTTCATTTTAATAGCTA	1550
QY	1502	AACAAACAAAGAGATCCAGAGTTTCAGAGCTATAATAAGAGTGAATTAATGATATA	1561
DB	1551	AACAAACAAAGAGATCCAGAGTTTCAGAGCTATAATAAGAGTGAATTAATGATATA	1610
QY	1562	GTTGACAAATGATATATCTCAGAAACCATGGCATCAATGACCAAAAAAATCCAGACA	1621
DB	1611	GTTGACAAATGATATATCTCAGAAACCATGGCATCAATGACCAAAAAAATCCAGACA	1670
QY	1622	CAGAAATATCGTACCATATGCTGCAATTTATGTGATTTCTAGCATTTGATTTGTCCAAC	1681
DB	1671	CAGAAATATCGTACCATATGCTGCAATTTATGTGATTTCTAGCATTTGATTTGTCCAAC	1730
QY	1682	ATAGTAGCAGTAGCTCACTGAGCTATTAATTAAGTGAATTAATTAAGTGAAGAA	1741
DB	1731	ATAGTAGCAGTAGCTCACTGAGCTATTAATTAAGTGAATTAATTAAGTGAAGAA	1790
QY	1742	TAAAAATTTAGCTCTTTCAGTGGTTAGCCACATGTAATAGTGGCTACCATCAGACG	1801
DB	1791	TAAAAATTTAGCTCTTTCAGTGGTTAGCCACATGTAATAGTGGCTACCATCAGACG	1850
QY	1802	GTGCAATATAGAAATTTCTTTTATAACAGAAAGTTCTATTGGAAAAAATTTCTTAGA	1861
DB	1851	GTGCAATATAGAAATTTCTTTTATAACAGAAAGTTCTATTGGAAAAAATTTCTTAGA	1910
QY	1862	AAATATACATAAATCTATAAAAAAAGAGAGTCAAGTGTGATTTGTTAGGCGAGGGT	1921
DB	1911	AAATATACATAAATCTATAAAAAAAGAGAGTCAAGTGTGATTTGTTAGGCGAGGGT	1970

Qy	1922	GAGGGAGATCGAATTGCAAAAGTGGTATGAGGAAAGTTTGGGGTAATAGGGTTGTTGGAA	1981
Db	1971		
Qy	1982	TCTTTGATGGGATGAGGGCTACTCGGTGCTTAATGTGTCACTCCTCCTCAGACTGCAACACATT	2041
Db	2031	TCTTGATGGGATGAGGGCTACTCGGTGCTTAATGTGTCACTCCTCCTCAGACTGCAACACATT	2090
Qy	2042	GGAAATTTGGCGAAATTCATCTGTATGTAAATATATACCTCATAAAGTAATCTCTAAAGAGTCAA	2101
Db	2091		
Qy	2102	GTGTTTTGTGGAAATTAATTTTTTAATCAGTTTGCAAATCTTATTATGAGATGATTTTTGCCAA	2161
Db	2151	GTGTTTTGTGGAAATTAATTTTTTAATCAGTTTGCAAATCTTATTATGAGATGATTTTTGCCAA	2210
Qy	2162	ATACATAAAACATGTTATTTCATTCATTAGGTGCAATATTTTTTGTCTAGCTCTCGTAACACACA	2221
Db	2211	ATACATAAAACATGTTATTTCATTCATTAGGTGCAATATTTTTTGTCTAGCTCTCGTAACACACA	2270
Qy	2222	GAGATGAATTTAGAAATAGCAAGCCTGCCCTCAAGCTGTTTCAATCCAGTACAGAGATGA	2281
Db	2271		
Qy	2282	GTCTATTCAAAAAATAGCTAGACTCCAGGAAGAAAGTTATAGGTGACCTTACACAAAAAAG	2341
Db	2331	GTCTATTCAAAAAATAGCTAGACTCCAGGAAGAAAGTTATAGGTGACCTTACACAAAAAAG	2390
Qy	2342	TGCAGATATAATTTATGTPAGCACAGTAGAAGTGGGAAAGTTCTTTTATGTGGAAAAAAG	2401
Db	2391	TGCAGATATAATTTATGTPAGCACAGTAGAAGTGGGAAAGTTCTTTTATGTGGAAAAAAG	2450
Qy	2402	AGGAGAAATTTTTTGGTCTTTTGAAGGATGACGAGATGTGAATATGCGCAGATCGAGTTTTT	2461
Db	2451	AGGAGAAATTTTTTGGTCTTTTGAAGGATGACGAGATGTGAATATGCGCAGATCGAGTTTTT	2510
Qy	2462	AAAACATTTCTGGTGGAGGGCAGAAATATGATCCAAGGCCACAGAGCAACACAGAAAAATAT	2521
Db	2511	AAAACATTTCTGGTGGAGGGCAGAAATATGATCCAAGGCCACAGAGCAACACAGAAAAATAT	2570
Qy	2522	GCAACCTAGAGGAAAGTGCAATGAAGGGAGCAGTTGTAAAATAATTTTCATGAATGTGAAG	2581
Db	2571	GCAACCTAGAGGAAAGTGCAATGAAGGGAGCAGTTGTAAAATAATTTTCATGAATGTGAAG	2630
Qy	2582	TGAGAAGAAATTTGTATCATAGACACTGAGTTTGGCAGAGTGCATGTTCTTGCGTCTCTAG	2641
Db	2631	TGAGAAGAAATTTGTATCATAGACACTGAGTTTGGCAGAGTGCATGTTCTTGCGTCTCTAG	2690
Qy	2642	GAGTCAAGAAAGAACAAAGTGTCCTTCTCTACGTTATGCTCAGTGGTCCCAAGTCCAAA	2701
Db	2691	GAGTCAAGAAAGAACAAAGTGTCCTTCTCTACGTTATGCTCAGTGGTCCCAAGTCCAAA	2750
Qy	2702	ACACCTTTCTCTTAAAGTACTTTCTTCTCCCTCCCATCAAAATCTAAAGTCTTTCACAA	2761
Db	2751	ACACCTTTCTCTTAAAGTACTTTCTTCTCCCTCCCATCAAAATCTAAAGTCTTTCACAA	2810
Qy	2762	ACATCATTTAAACACGGCAGTTCATGGTCAGAAAGGCAATTTGCTTTTCTTAGACTTCTATG	2821
Db	2811	ACATCATTTAAACACGGCAGTTCATGGTCAGAAAGGCAATTTGCTTTTCTTAGACTTCTATG	2870
Qy	2822	TACGTTATTATTTACAAATTTCTCGCTTAAAGACCTCTAAAGTCTTGGAAGTTTCCACC	2881
Db	2871	TACGTTATTATTTACAAATTTCTCGCTTAAAGACCTCTAAAGTCTTGGAAGTTTCCACC	2930
Qy	2882	TTGCACATCAAAAGATATAAATTCATGCAATTTGTATAGTAACCTTAGTCCCTTAAGAGAATA	2941
Db	2931	TTGCACATCAAAAGATATAAATTCATGCAATTTGTATAGTAACCTTAGTCCCTTAAGAGAATA	2990
Qy	2942	AGGATGAACCTATAAATATAAGAAGTAATTTATGGTAAATTTATTAATATGATTTGCCACTTATTT	3001
Db	2991	AGGATGAACCTATAAATATAAGAAGTAATTTATGGTAAATTTATTAATATGATTTGCCACTTATTT	3050

Qy	3002	TTCACTTGATCGTGATGGTTGCATCTACTGGTGTCTCTGTTGCAATCTCTAGAGAGTTTCG	3066
Db	3051	TTCACTTGATCGTGATGGTTGCATCTACTGGTGTCTCTGTTGCAATCTCTAGAGAGTTTCG	3110
Qy	3062	CTCTTTTCTCGGGTCAAACTCTCGCCATTTATTTTCATAAATGCAATAGAGAGCCAACTCTTT	3121
Db	3111	CTCTTTTCTCGGGTCAAACTCTCGCCATTTATTTTCATAAATGCAATAGAGAGCCAACTCTTT	3170
Qy	3122	TTCCATAATTACTTATTTTAAAAATTTGTTGGCATTTAAATTTCTGTTCTCTTTAGCTTAGTAA	3181
Db	3171	TTCCATAATTACTTATTTTAAAAATTTGTTGGCATTTAAATTTCTGTTCTCTTTAGCTTAGTAA	3230
Qy	3182	CTTTAGGATTTTAAATPACAACTATTGAAATCATACACATACGTTTTAAATGATTAATTTT	3241
Db	3231	CTTTAGGATTTTAAATPACAACTATTGAAATCATACACATACGTTTTAAATGATTAATTTT	3290
Qy	3242	AAATACGTTAGGCTATAAACCTTTTAAATTTTAAAAAAATAGATGAGTGTGGTGGCTC	3301
Db	3291	AAATACGTTAGGCTATAAACCTTTTAAATTTTAAAAAAATAGATGAGTGTGGTGGCTC	3350
Qy	3302	ATGCGCTGTAATCCCAACACTTTTGGGAAGCGGGTCCGGAGGATAGCTTGAGTCCAGCAGT	3361
Db	3351	ATGCGCTGTAATCCCAACACTTTTGGGAAGCGGGTCCGGAGGATAGCTTGAGTCCAGCAGT	3410
Qy	3362	TTGAGACCACTGAGGGCAACACAGAAAGCCCAATATCTAAAAAAACAAAACAAA	3421
Db	3411	TTGAGACCACTGAGGGCAACACAGAAAGCCCAATATCTAAAAAAACAAAACAAA	3470
Qy	3422	ATTACCTGGGTATGGTTGTGCTCCACCTGTAGTCCAAAGTACACAGGAAGCTGAGGCGAGAA	3481
Db	3471	ATTACCTGGGTATGGTTGTGCTCCACCTGTAGTCCAAAGTACACAGGAAGCTGAGGCGAGAA	3530
Qy	3482	GGATCACTTGAGCGCCAGGAGTTGAGGCTGCAGCTGATCCATGAACCGCTGCTACACTCA	3541
Db	3531	GGATCACTTGAGCGCCAGGAGTTGAGGCTGCAGCTGATCCATGAACCGCTGCTACACTCA	3590
Qy	3542	GTCTGGGTGACGTGCAAGAAGCTGTCTCAAAAATATAAATAAATAAATAAATAAATAAATAA	3601
Db	3591	GTCTGGGTGACGTGCAAGAAGCTGTCTCAAAAATATAAATAAATAAATAAATAAATAAATAA	3650
Qy	3602	AAAAACAAAAATTAATTTAAATTTTAAAAACACACACACACTAGAGATGTTTTGCAAAATTGAT	3661
Db	3651	AAAAACAAAAATTAATTTAAATTTTAAAAACACACACACACTAGAGATGTTTTGCAAAATTGAT	3710
Qy	3662	TATTTGGGAGTCTATATCCCTGGAAGTTAATTTAAAAATATTTAAAGAGAGTTCTTCCCTCAT	3721
Db	3711	TATTTGGGAGTCTATATCCCTGGAAGTTAATTTAAAAATATTTAAAGAGAGTTCTTCCCTCAT	3770
Qy	3722	TTCTTAGAGACGTGGAATTTGAAATATCAGAGCTAGAAAGAACACTAGGGGTGCGCACTC	3781
Db	3771	TTCTTAGAGACGTGGAATTTGAAATATCAGAGCTAGAAAGAACACTAGGGGTGCGCACTC	3830
Qy	3782	CAAAAGTGTGTCNAGGACACAGCAGCATCAAGTAAACCTGGGAACGTGTAGAAATGAGA	3841
Db	3831	CAAAAGTGTGTCNAGGACACAGCAGCATCAAGTAAACCTGGGAACGTGTAGAAATGAGA	3890
Qy	3842	GTCTTTAGGCTTCAACCCAGACCTTACTGAAACAGAAATCTGCAATTAACAAAGATTTCTAGGTG	3901
Db	3891	GTCTTTAGGCTTCAACCCAGACCTTACTGAAACAGAAATCTGCAATTAACAAAGATTTCTAGGTG	3950
Qy	3902	CCTCAGCGGCACATTAATACTGAGAAAGCTGTGCACTAGAAATCTTCACTCCACCTTTTCA	3961
Db	3951	CCTCAGCGGCACATTAATACTGAGAAAGCTGTGCACTAGAAATCTTCACTCCACCTTTTCA	4010
Qy	3962	TTATAAATGGAATCACTTGGGCTGTGTGTACAGGAAATTCATTTTTTAAATTTTCAGAAC	4021
Db	4011	TTATAAATGGAATCACTTGGGCTGTGTGTACAGGAAATTCATTTTTTAAATTTTCAGAAC	4070
Qy	4022	CTTCTATTTAGGTCATCTATATTTGCTTAATAGCAGGGAAGAAAGCCAAACCTCTTTAACTG	4081
Db	4071	CTTCTATTTAGGTCATCTATATTTGCTTAATAGCAGGGAAGAAAGCCAAACCTCTTTAACTG	4130
Qy	4082	CAATTAACAAATCTATAATTAATTAGTTAAAGCAATCTTCCCTTTTAAAGTTTTTACATTTTGT	4141

Db	4131	CAATTAA	CAAAATCTATAATTAATTAAGTTAAGCAATCTTCCCTTTAAGTTTACATTTTGT	4190
Qy	4142	GGAGCAAGCTGTTTGAATTCGGCTGGGCTCAGGCCGCGCTGTTTGAAATTCACAAATTC	4201	
Db	4191	GGAGCAAGCTGTTTGAATTCGGCTGGGCTCAGGCCGCGCTGTTTGAAATTCACAAATTC	4250	
Qy	4202	ACAGATGTTAGCCGCTCTCCGGCTTAAGTAAAGGAAGAGAAATGTCAGTTTAAATAGCTT	4261	
Db	4251	ACAGATGTTAGCCGCTCTCCGGCTTAAGTAAAGGAAGAGAAATGTCAGTTTAAATAGCTT	4310	
Qy	4262	CTCCCTTCCATCTCGGCTGAAGCAACAAATAAATAATTTTTATGAACACATTTTGAGTT	4321	
Db	4311	CTCCCTTCCATCTCGGCTGAAGCAACAAATAAATAATTTTTATGAACACATTTTGAGTT	4370	
Qy	4322	AGATTTACTTACAGGGAATGTCAAATTTCTCTGAAGGCTTTAGATTGTCACAACT	4381	
Db	4371	AGATTTACTTACAGGGAATGTCAAATTTCTCTGAAGGCTTTAGATTGTCACAACT	4430	
Qy	4382	TTGACATCTACTGATGTCACTATTTACAGTGTCTCTGTGACTAGGGGGTGAAGGAA	4441	
Db	4431	TTGACATCTACTGATGTCACTATTTACAGTGTCTCTGTGACTAGGGGGTGAAGGAA	4490	
Qy	4442	GATGTGAATCACCATGTTAGTGAACGTTAGATACACAGAGTGGTTTTTTTCCCGCTGT	4501	
Db	4491	GATGTGAATCACCATGTTAGTGAACGTTAGATACACAGAGTGGTTTTTTTCCCGCTGT	4550	
Qy	4502	TGGAGTCTATCTTAACCTGAGCTTCTGAATCATATTTCAATTTCCAAATCCACAAA	4561	
Db	4551	TGGAGTCTATCTTAACCTGAGCTTCTGAATCATATTTCAATTTCCAAATCCACAAA	4610	
Qy	4562	CCAGGATAAGTTTACAGCCCATATTCAGAAAGGAATAAATAATTTTGTGTAGACTTT	4621	
Db	4611	CCAGGATAAGTTTACAGCCCATATTCAGAAAGGAATAAATAATTTTGTGTAGACTTT	4670	
Qy	4622	CCTGATATTACACTGATTTCCGGAATATATGAACAAATTTTATGTTTCTTCGAAAGTAGG	4681	
Db	4671	CCTGATATTACACTGATTTCCGGAATATATGAACAAATTTTATGTTTCTTCGAAAGTAGG	4730	
Qy	4682	TCAAGTCAAAAGCAAAACCAAAAACAGCAAAACTGTGAAGACATAAAGATAGAGTGGAC	4741	
Db	4731	TCAAGTCAAAAGCAAAACCAAAAACAGCAAAACTGTGAAGACATAAAGATAGAGTGGAC	4790	
Qy	4742	CGACTGAGAGATTAATAATAACTAGAAATATTTTATTAACAGGCAATTTTCAAAATAATTCG	4801	
Db	4791	CGACTGAGAGATTAATAATAACTAGAAATATTTTATTAACAGGCAATTTTCAAAATAATTCG	4850	
Qy	4802	TGCACCTTCAGAAATTTTACAATAATATATTTTCCAAATTTTAATATCTTTTAAGAAAAT	4861	
Db	4851	TGCACCTTCAGAAATTTTACAATAATATATTTTCCAAATTTTAATATCTTTTAAGAAAAT	4910	
Qy	4862	TACTATATTATATGTAAGTACATGTGTCATGTGTTTGAAGTAGGATATTTAACTCAATAAA	4921	
Db	4911	TACTATATTATATGTAAGTACATGTGTCATGTGTTTGAAGTAGGATATTTAACTCAATAAA	4970	
Qy	4922	GGTATTTTCTTTATTCGGGTGAGCAAGCTTCTTAAGGGGATGTCAAGGGATATCTC	4981	
Db	4971	GGTATTTTCTTTATTCGGGTGAGCAAGCTTCTTAAGGGGATGTCAAGGGATATCTC	5030	
Qy	4982	TTTCTCTTAGCTGAGAGGAAGTAGTCTTCAAGTTTAAATATAATCAAGGAAATTTCCCTG	5041	
Db	5031	TTTCTCTTAGCTGAGAGGAAGTAGTCTTCAAGTTTAAATATAATCAAGGAAATTTCCCTG	5090	
Qy	5042	TCCTTGCTATTGAGATTGTGACCAACAGCGGTTGGCTGAAGGAACTGAAGGGC	5101	
Db	5091	TCCTTGCTATTGAGATTGTGACCAACAGCGGTTGGCTGAAGGAACTGAAGGGC	5150	
Qy	5102	GGGAGGGAGGGAATAGATGAAACAAACAAACAAACAACTTCCCTAAGCAGCTCTA	5161	
Db	5151	GGGAGGGAGGGAATAGATGAAACAAACAAACAAACAACTTCCCTAAGCAGCTCTA	5210	
Qy	5162	CAAAACATTTTAGCCCAAGAAATAGTCAAGAAATCTTCAAAATCAAAACCAAGTATCCAGAT	5221	

Db	5211	CAAAACATTTTAGCCCCAGAAATAGTCAAGAAATCTTCAAAATCAAAACAGTATCCAGAT	5270	
Qy	5222	ACAAGCAAGCTGTTTATCTAGCTGGAGCAGGCTGACACTCATCAGCTCAGTTTCAGTTTACAA	5281	
Db	5271	ACAAGCAAGCTGTTTATCTAGCTGGAGCAGGCTGACACTCATCAGCTCAGTTTCAGTTTACAA	5330	
Qy	5282	AAGTCCAGGCTGCTGAAATTTAAACTCTGATGCCATTTCATGCCAGCTTCCAATCACGACAG	5341	
Db	5331	AAGTCCAGGCTGCTGAAATTTAAACTCTGATGCCATTTCATGCCAGCTTCCAATCACGACAG	5390	
Qy	5342	AGATCAAGATTCAGAGATGCCCTCCAGCTCCAAAATTCGCAACAAACAGTGTGGCTACTAT	5401	
Db	5391	AGATCAAGATTCAGAGATGCCCTCCAGCTCCAAAATTCGCAACAAACAGTGTGGCTACTAT	5450	
Qy	5402	ACGTCAAGCACTCTGAAGCCGTGAGAGGCGGGAACAAACAGTAGAGGATGCCAG	5461	
Db	5451	ACGTCAAGCACTCTGAAGCCGTGAGAGGCGGGAACAAACAGTAGAGGATGCCAG	5510	
Qy	5462	CTGGTAAGAAATCGAGTGTGTTTATGAAGTTTATGATGAATTTCAATTGCTAAAAAT	5521	
Db	5511	CTGGTAAGAAATCGAGTGTGTTTATGAAGTTTATGATGAATTTCAATTGCTAAAAAT	5570	
Qy	5522	CAAGAAACGCTCGGCTCTTTTGCAAAATATGATGAAGGAGAGAAAGTCCCTAAACTTCTAT	5581	
Db	5571	CAAGAAACGCTCGGCTCTTTTGCAAAATATGATGAAGGAGAGAAAGTCCCTAAACTTCTAT	5630	
Qy	5582	GTCTGATAGCATTTGACCTATTTGCTTTTAGCTCCCGCTTTATATCTATATATATACACA	5641	
Db	5631	GTCTGATAGCATTTGACCTATTTGCTTTTAGCTCCCGCTTTATATCTATATATATACACA	5690	
Qy	5642	GGTATTTGTGTATATTTTATATAAATTTGTTCTCCGT	5676	
Db	5691	GGTATTTGTGTATATTTTATATAAATTTGTTCTCCGT	5725	
RESULT 6				
ADO48537/c				
ID	ADO48537 standard; DNA; 58326 BP.			
XX				
AC	ADO48537;			
XX				
DT	12-AUG-2004 (first entry)			
XX				
DE	Human mannose receptor C type 2 (MRC2) genomic DNA sequence.			
XX				
XX	human; melanoma; single nucleotide polymorphism; SNP; neuropilin 1; NRPL1;			
KW	mannose receptor C type 2; MRC2; gene; d6; ENDO180.			
XX	Homo sapiens.			
OS				
XX				
PH	Key Location/Qualifiers			
FT	variation 43763			
FT	/*tag= a			
FT	/note= "Single nucleotide polymorphism"			
XX				
PN	WO2004044163-A2.			
XX				
PD	27-MAY-2004.			
XX				
PF	06-NOV-2003; 2003WO-US035876.			
XX				
PR	06-NOV-2002; 2002US-0424475P.			
PR	23-JUL-2003; 2003US-0489703P.			
XX				
PA	(SEQU-) SEQUENOM INC.			
XX				
PI	Roth RB, Nelson MR, Braun A, Kammerer SM;			
XX				
DR	WPI; 2004-411720/38.			
XX				
PT	Identifying a subject at risk of melanoma, useful for treating melanoma, comprises detecting the presence or absence of one or more polymorphic variations associated with melanoma in a nucleic acid sample from a			

```
FT subject.
XX Claim 23; SEQ ID NO 3; 176pp; English.
XX
CC The invention comprises a method for identifying a subject at risk of
CC melanoma. The invention involves detecting the presence or absence of one
CC or more polymorphic variations associated with melanoma in the neuropilin
CC 1 (NRP1) or mannose receptor C type 2 (MRC2) genes. The method of the
CC invention is useful for identifying subjects at risk and treating
CC melanoma. The present nucleic acid represents the genomic DNA sequence
CC for human MRC2.
XX
SQ Sequence 58326 BP; 13666 A; 14813 C; 15614 G; 14232 T; 0 U; 1 Other;
Query Match 3.5%; Score 196.2; DB 12; Length 58326;
Best Local Similarity 70.0%; Pred. No. 5.2e-27;
Matches 278; Conservative 0; Mismatches 118; Indels 1; Gaps 1;
Qy 3262 CTTTAAATTTTAAATAATAGATGAGTGTGGTCTCATGCTGTAAATCCCAACACT 3321
Db 18059 CTATAAAGTTTTCACAGGGAAGCCAGGTGCGTGGCTCACACCTGTATCCAGCACT 18000
Qy 3322 TTGGGAAGCGGGTCGGGAGATAGCTTGTAGTCAGCAGTTTGAGACAGTCAGGGCAAC 3381
Db 17999 TTGGGAAGCTGAGCAGGTGATTTGCTTGTAGCTCAGGAGTTTGAGACCAGCTGGGCAAC 17940
Qy 3382 ACACGACGCCCATATCTAATAAAACAAACAAACAAATACCTGGGTATGGTGTG 3441
Db 17939 ATGCAAAAACCCATCTCTACATAAAATACAGAAAAAAATAGCTGGTGGTG 17880
Qy 3442 CTACCTGTAGTCCAAAGTCACAGGAAGCTGAGGCAGAGATCACTTGTAGCCCGAGG 3501
Db 17879 CACACCTGTGGTCTGCTTACTCAGGATCTGAGATGGGAGTGTCTTGTAGCCCAAG 17820
Qy 3502 GTTGAGCTGCAGTGATCCATGAACGGCTGCTACACT-CAGTCTGGTGACAGTGAAG 3560
Db 17819 ATTGAGCTGCAGTGAGCAGCAAGATCATGCTACTTCCAGCTGGGCAACAGCAAG 17760
Qy 3561 AAGCTGTCTCAAAATATATAATATAAATAAATAAATACCTTTTAAACAAATAATTA 3620
Db 17759 ACTCCATCTCAAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 17700
Qy 3621 ATTTTAAACACACACACTAGAGATGTTTGCAAT 3657
Db 17699 TATAATATATACACACACATATATGTGTATAT 17663
RESULT 7
ADP75180/c
ID ADP75180 standard; DNA; 304905 BP.
XX
AC ADP75180;
XX
DT 12-AUG-2004 (first entry)
XX
DE Human Endophilin 2 gene.
XX
KW Human; chromosome 19; ds; gene; ADAM19; Endophilin 1; Endophilin 2; NRG2;
KW ADAMTS2; a disintegrin and metalloprotease; neuroregulin 2; SNP;
KW single nucleotide polymorphism;
KW a disintegrin and metalloprotease with thrombospondin type1 motif 2;
KW asthma; atopy; obesity; inflammatory bowel disease; respiratory disorder.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT variation replace(276365,T)
FT /tag= a
FT /standard name= "Single nucleotide polymorphism"
FT variation replace(276366,G)
FT /tag= b
FT /standard name= "Single nucleotide polymorphism"
FT variation replace(278194..278197,CT)
```

```
FT /tag= c
FT /standard name= "Single nucleotide polymorphism"
FT variation replace(278474,A)
FT /tag= d
FT /standard name= "Single nucleotide polymorphism"
FT variation replace(278990,C)
FT /tag= e
FT /standard name= "Single nucleotide polymorphism"
FT variation replace(279072,G)
FT /tag= f
FT /standard name= "Single nucleotide polymorphism"
FT variation replace(280049,T)
FT /tag= g
FT /standard name= "Single nucleotide polymorphism"
FT variation replace(280107,A)
FT /tag= h
FT /standard name= "Single nucleotide polymorphism"
FT variation replace(280124,T)
FT /tag= i
FT /standard name= "Single nucleotide polymorphism"
PN WO2003031594-A2.
XX
XX 17-APR-2003.
XX
XX 11-OCT-2002; 2002WO-US032700.
XX
XX 11-OCT-2001; 2001US-0328424P.
XX
XX (GENO-) GENOME THERAPEUTICS CORP.
XX
XX Keith T, Little RD, Van Eerdewegh P, Dupuis J, Del Mastro RG;
XX Allen K;
XX
XX WPI; 2003-381712/36.
XX
XX New isolated nucleic acid or alternate splice variant, useful for
XX diagnosing and treating a disintegrin and metalloprotease (ADAM) or
XX interactor gene-associated disorder, e.g. asthma, atopy, obesity or
XX inflammatory bowel disease.
XX
XX Claim 2; SEQ ID NO 1; 338pp; English.
XX
XX The invention relates to an isolated nucleic acid or alternate splice
XX variant comprising a nucleotide sequence containing at least one of the
XX single nucleotide polymorphisms given in the specification, a nucleotide
XX sequence having at least 15 contiguous nucleotides of them, or
XX complements of them. The genes are ADAM19 (a disintegrin and
XX metalloprotease 19, also known as gene 845), NRG2 (neuroregulin 2, also
XX known as gene 847), endophilin 1 (also known as gene 874), endophilin 2
XX (also known as gene 803) and ADAMTS2 (a disintegrin and metalloprotease
XX with thrombospondin type1 motif 2, also known as gene 962). Also included
XX are a vector comprising the isolated nucleic acid (or alternate splice
XX variant), a host cell containing the vector, an isolated polypeptide
XX encoded by the novel nucleic acid (or alternate splice variant), an
XX antibody or antibody fragment that binds to the polypeptide,
XX pharmaceutical compositions (comprising the nucleic acid or alternate
XX splice variant, vector, polypeptide or antibody, and a carrier,
XX excipient or diluent), a kit for detecting a disintegrin and
XX metalloprotease (ADAM) gene nucleotide sequence (comprising the isolated
XX nucleic acid or alternate splice variant, antibody or antibody fragment,
XX and at least one component to detect the hybridisation of the variant or
XX the binding of the antibody to an ADAM gene amino acid sequence), a kit
XX for detecting an interactor gene amino acid sequence (comprising the
XX antibody or antibody fragment, and at least one component to detect the
XX binding of the antibody to the interactor gene amino acid sequence),
XX diagnosing an ADAM or interactor gene-associated disorder or a
XX respiratory disorder in a human subject, determining an ADAM or
XX interactor gene pharmacogenetic profile in a human subject, identifying
XX an orthologue of a human ADAM or interactor gene, treating an ADAM or
XX interactor gene-associated disorder (or a respiratory disorder) by
XX administering the pharmaceutical composition, a transgenic mouse (whose
XX genome comprises an introduced null mutation in an endogenous gene that
```

is orthologous to a human ADAM gene), making a homozygous transgenic knockout mouse, forming a crystal of the isolated polypeptide, a cell line comprising the isolated nucleic acid or alternate splice variant, a biochip comprising the isolated nucleic acid or alternate splice variant, an isolated nucleic acid probe or primer comprising at least 8 contiguous nucleotides of the nucleic acid, an isolated antisense nucleic acid, identifying an ADAM or interactor gene ligand and an isolated nucleic acid variant of Gene 803, 845, 847, 874 or 962. The nucleic acid or alternate splice variants, methods, kits and antibody/antibody fragment are useful for diagnosing and treating an ADAM or interactor gene-associated disorder, e.g. asthma, atopy, obesity or inflammatory bowel disease. The present sequence is a gene (or gene fragment) for one of the above mentioned genes.

Sequence 304905 BP; 71522 A; 82757 C; 84652 G; 65974 T; 0 U; 0 Other;
Query Match 3.5%; Score 196; DB 11; Length 304905;
Best Local Similarity 73.3%; Pred. No. 7.1e-27;
Matches 264; Conservative 0; Mismatches 95; Indels 1; Gaps 1;

3279 AAAATAGATGAGTGTGGTCTCATGCTCTAATCCCAACACTTTGGGAAGCGGGTCGG 3338
156005 AATCTCGTGGGTGCGGTAGCTACACCTGTAATCCAGCAGCTTGGGAGGCCAAGGTGG 155946

3339 GAGGATAGCTTGACTCCAGCAGTTTGAGACAGCTCAGGCGCAACACAGCAAGACCCCATAT 3398
155945 GAGGATTTGCTTAAGCCAGAGTTTGAGACAGCGCTGGGCAACATGTTGAACCCCATGT 155886

3399 CTAATAAATAAACAACAAATTAACCTGGGTGATGTTGTCTCACTCTGTAGTCCAAG 3458
155885 CTAATAAATAAATAAACAACAACTGTAGCAGGTGTGGTGCACACCTGTAGTTTCAG 155826

3459 CTACACAGGAAGCTGAGCGAAGGATCACTTGAGCCAGAGGTTGAGGCTGCGAGTAT 3518
155825 CTGCTCAGGAGGCTGAGGTGGGAGATCACCCTGAGCCAGGAAAGTTTGAGGCTGCGAGT 155766

3519 CCATGAACGCGCTGTACACT -CAGTCTGGGTGACAGTGAAGAAGCTGTCTCAAAAAATA 3577
155765 TCGTATCGGCGCCACTGCACTCCAGCCTGGGTGACAGATGAACCCCGTCTCAAAAAATA 155706

3578 ATAATAAATAAATAAATTAATTTAAAAAACAATAAATTAATTTAAAAACAACA 3637
155705 AAAAAAATAAATAAACAACAACCAAGCAACTCAATCTCTCTCTCTTCTTAAAAAAGCA 155646

RESULT 8
ABX56555/c
ID ABX56555 standard; DNA; 118384 BP.

AC ABX56555;

XX 20-FEB-2003 (first entry)

DE Human autoimmune disease related protein PAT1 gene region #1.

XX PAT1; human; autoimmune disease; psoriasis; type I diabetes;
XX rheumatoid arthritis; cation-chloride transport; gene therapy; gene; ds.

OS Homo sapiens.

XX WO200280842-A2.

XX 17-OCT-2002.

XX 04-APR-2002; 2002WO-SE000669.

XX 05-APR-2001; 2001GB-00008587.

XX 05-APR-2001; 2001GB-00008589.

XX 03-APR-2001; 2001GB-00008590.

XX (ESTE-) ESTECON AB.

XX Swanbeck G, Purvis I, Dykes C, Hewett D;

XX WPI; 2003-058462/05.
DR New PAT1 genes and polypeptides, useful for diagnosing the presence of,
XX or a susceptibility to, an autoimmune disease (e.g. psoriasis, type I
PT diabetes or rheumatoid arthritis), or screening agents for treating the
PT autoimmune diseases.
XX
PS Claim 18; Page 74-104; 114pp; English.

XX The invention describes an isolated polypeptide expressed from the PAT1
CC (gene) region, its homologue, variant or fragment. The polypeptide or
CC polynucleotide encoding it are useful for diagnosing the presence of, or
CC a susceptibility to, an autoimmune disease such as psoriasis, type I
CC diabetes or rheumatoid arthritis. The polynucleotide or polypeptide are
CC also useful for screening its modulators. The agent that modulates PAT1,
CC or the component that regulates cation-chloride transport, is affected by
CC action-chloride transport, regulates PAT1 or is affected by PAT1 is
CC useful for treating, or in the manufacture of a medicament for preventing
CC or treating, an autoimmune disease. This sequence represents a
CC polynucleotide from the human PAT1 gene
XX

Sequence 118384 BP; 31464 A; 26095 C; 27139 G; 33686 T; 0 U; 0 Other;
Query Match 3.4%; Score 193.4; DB 10; Length 118384;
Best Local Similarity 68.8%; Pred. No. 1.9e-26;
Matches 295; Conservative 0; Mismatches 131; Indels 3; Gaps 2;

3276 AAAAAAATAGATGAGTGTGGTCTCATGCTGTATCCCAACACTTTGGGAAGCGGGT 3335
114643 AATAAAAGGCTGAGTGTGGTCTCACACCTGTATCCAGCAGCTTGGGAAGCCAAG 114584

3336 CGGGAGGATAGCTTGAGTCCAGCAGTTTGAGACCACTCAGGCGCAACACAGCAAGACCCCA 3395
114583 CAGGTGGATTACTTGAGGTCAGGAGTTCGAGACCACTGGCCCAACATGGTGAACCCCTG 114524

3396 TATCTAAAAAACAACAAAAAATAATTTACCTGGGTATGGTTGTCTCACCTGTAGTCC 3455
114523 TCTCTATTAAA -AACAAAAACAATAATTTAGCTGGGTGTGTACACGCTGTATATCC 114466

3456 AAGCTACACAGGAAGCTGAGGCGAGAGATCACTTGAGCCAGGAGGTTGAGGCTGCAGT 3515
114465 CAGCTACTTTGGGAAGCTGAGGTGGGAGGATCACTTGAACCCAGGAGCGAGGTTGTAGC 114406

3516 GATCCATGAACGGCTGTCTACT -CAGTCTGGGTGACAGTGCAGCAAGCTGTCTCAAAA 3574
114405 GAGCCAAAGATCATGCCACTGCATCCATCTCTGGGTATAGACCAAGCACTCCATCTCANA 114346

3575 ATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 3634
114345 ATGAAAAGAAAACAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 114286

3635 ACACACTAGAGATGTTTGCATAATTTGTTGGGAGTCTATATCCCTCGAAGTTAATTT 3694
114285 CCAGTAACATGTTGTTTATTATCAAGTATTATGTACTGTATATAATTTGATTAAATGT 114226

3695 AAAATATTT 3703

114225 GCTATACTT 114217

RESULT 9
AAC74312
ID AAC74312 standard; cDNA; 1384 BP.

XX AAC74312;

XX 02-FEB-2001 (first entry)

XX Human secreted protein gene 33 SEQ ID NO:43.

XX Human; secreted protein; diagnosis; immunosuppressive; antiarthritic;
XX antirheumatic; antiproliferative; cytostatic; cardiac; vasotropic;

[illegible]

Matches 260; Conservative 0; Mismatches 101; Indels 1; Gaps 1;
QY 3263 TTTTAAATTTTAAATAATAGATGAGTGTGGGCTCATGCTGTAAATCCCAACACTT 3322
Db 26435 TTTTATAAATGTGACAGTGGCTGAGTGAGTGGCTCATGCTGTAAATCCCAACACTT 26494
QY 3323 TGGGAAGCCGGGTGCGGAGGATAGCTTTGAGTCCAGCAGTTTGAGACCAGTCAGGGCAACA 3382
Db 26495 TGGGAGCTGAGCGAGGGGATTCCTTGAGCCCATGAGTTTGAGACCAGCTGGCCAAACA 26554
QY 3383 CAGCAAGACCCCATATTTAAATAAACAACAAACAAATTAACCTGGGTATGTTGTC 3442
Db 26555 GGGTAAATCCCATCTCAACAAAAATACACACAAAAATTAAGCTGGCATGGTGTGT 26614
QY 3443 TCACCTGTAGTCCAGCTACACAGGAGCTGAGCAGAGGATCATTGAGCCAGGAGG 3502
Db 26615 ACATCTGTGGTCCAGCTACTCAGGAGCGGAGTGGGAGGATCACCCTGAGCCCGAGG 26674
QY 3503 TTGAGGCTGACAGTGCATGCAAGCCGCTGTACACT-CAGTCTGGGTGACAGTGAAGA 3561
Db 26675 TCTAGACTGACGAGCTGTGATTTGTGCCACTGCCTCCAGTCTGGGCAACAGAGTGAGA 26734
QY 3562 AGCTGTCTCAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 3621
Db 26735 CCCTGTCTAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 26794
QY 3622 TT 3623
Db 26795 AT 26796

RESULT 11
ABN99663
ID ABN99663 standard; DNA; 8133 BP.
AC ABN99663;
XX
DT 16-AUG-2002 (first entry)
XX
DE Human clusterin gene sequence 2.
XX
KW Human; antisense inhibition; antisense oligonucleotide; clusterin;
KW hypercholesterolaemia; cardiovascular disorder; ds;
KW hyperproliferative disorder; hyperlipidemic disorder.
XX
OS Homo sapiens.
XX
XX WO200222635-A1.
XX
PD 21-MAR-2002.
XX
PF 10-SEP-2001; 2001WO-US028235.
XX
PR 11-SEP-2000; 2000US-00659791.
XX
PA (TSIS-) ISIS PHARM INC.
XX
PI Monia BP, Freier SM;
XX
XX WPI; 2002-404805/43.
XX
PT Novel antisense compound targeted to nucleic acid molecule encoding
PT clusterin, useful for treating animal having disease associated with
PT clusterin such as hyperlipidemic disorder, cardiovascular disorder.
XX
PS Example 15; Page 93-98; 125pp; English.
XX
CC The invention comprises antisense oligonucleotides that are capable of
CC inhibiting expression of the human clusterin gene. The antisense
CC oligonucleotides of the invention are useful for inhibiting the
CC expression of clusterin in cells. The antisense oligonucleotides are also
CC useful for treating an animal with a disease or condition associated with
CC clusterin (e.g. hypercholesterolaemia; cardiovascular disorders;

CC hyperproliferative disorders; and hyperlipidemic disorders). The present
CC DNA sequence represents a human clusterin gene sequence
SQ Sequence 8133 BP; 1982 A; 2024 C; 2002 G; 2125 T; 0 U; 0 Other;
Query Match 3.3%; Score 187.6; DB 6; Length 8133;
Best Local Similarity 74.6%; Pred. No. 1.7e-25;
Matches 249; Conservative 0; Mismatches 84; Indels 1; Gaps 1;
QY 3264 TTTTAAATTTTAAATAATAGATGAGTGTGGTCTATGCTGTAAATCCCAACACTTT 3323
Db 1057 TTTTAAAGAGGAAAAAATGAGCTGGGCATGGTGGCTCATGCTGTGATGCCAGCACTTT 1116
QY 3324 GGAAGCCGGGTGCGGAGGATAGCTTTGAGTCCAGCAGTTTGAGACCAGTCAGGGCAACAC 3383
Db 1117 GAGAGCCAAAGGGGGAGGATCATTTGAGGCCAGGAGTTTGAGACCAGCTGCGCAAGAT 1176
QY 3384 AGCAAGACCCCATATCTAAAAAACAACAAAAATAAATAAATAAATAAATAAATAAATAA 3443
Db 1177 AGTGAGACCTCTCTACAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1236
QY 3444 CACCTGTAGTCCAGCTACACAGGAAGCTGAGCAGAAAGATCATTTGAGCCCGAGGAGT 3503
Db 1237 CACCTGTAGTCTCAGCTACTCGGGAGCTGAGCAAGAGAGTCACTGAGCCTGGAAGTT 1296
QY 3504 TGAGGCTGCAAGTATCCATGAACGCGCTGCTACACT-CAGTCTGGGTGACAGTGAAGA 3562
Db 1297 GGAGGCTGCAAGTATGATTTGACCATTCAGCTGCGGCAACAGAGTGAGAC 1356
QY 3563 GCTGTCTCAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 3596
Db 1357 CCTGTCTTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1390

RESULT 12
ADJ12659
ID ADJ12659 standard; DNA; 18861 BP.
XX
AC ADJ12659;
XX
DT 20-MAY-2004 (first entry)
XX
DE DNA fragment of a BAC clone that encodes a human secreted protein Seq513.
XX
KW human; secreted; cancer; haematopoietic disease; anaemia;
KW multiple myeloma; reproductive system disorder; prostatitis;
KW inguinal hernia; musculoskeletal disease; systemic lupus erythematosus;
KW gout; cardiovascular disease; arrhythmia; hypernatraemia; fetal disease;
KW fetal alcohol syndrome; Down's syndrome; excretory disease;
KW urinary incontinence; renal disorder; neural; sensory disease;
KW Alzheimer's disease; meningitis; respiratory disease; diabetes;
KW occupational lung disease; endocrine disease; diabetes;
KW glomerulonephritis; digestive disease; portal hypertension;
KW irritable bowel syndrome; epithelial disease; scleroderma;
KW epidermolysis bullosa; cytostatic; antianemic; antiarthritic;
KW antiasthmatic; anti-HIV; immunosuppressive; antiinflammatory;
KW antipsoriatic; antibacterial; osteopathic; dermatological; antitout;
KW immunomodulator; antiarrhythmic; cardiac; nootropic; antileptic;
KW nephrotropic; uropathic; neuroprotective; antiparkinsonian; tranquilizer;
KW antidiabetic; anabolic; hypertensive; vulnery; ds.
XX
OS Homo sapiens.
XX
PN US2004010132-A1.
XX
PD 15-JAN-2004.
XX
PF 30-OCT-2001; 2001US-00984429.
XX
PR 09-OCT-1997; 97US-0061463P.
PR 09-OCT-1997; 97US-0061527P.
PR 09-OCT-1997; 97US-0061529P.
PR 09-OCT-1997; 97US-0061532P.

D	b		5617	GGAGGCCGAGGTGGGAGATCACTTGAGCTCAGGAGTTTGGACACCAGCCTGAGCAACATA	5678
Q	y		3385	GCAAGACCCCATATCTAAAAAAACAACAACAATAAATTACTCGGATGTTGTCCTC	3444
D	b		5677	GCAAGACTTTCATCTCTACTAAAAGAAAAAGTTAAAAAAATTTAGCCAGCGCGTGGTGCAC	5736
Q	y		3445	ACCTGTAGTCCAGCTACACAGGAAGCTGAGGCAGAGAAGATCACTTTGAGCCCAGGAGGTT	3504
D	b		5737	ACCTGTAGTCCCCAGCTACTTCAGAAAGCTGAGGCAGGAGGATCACTTTGAGTCTCGGAGGTT	5796
Q	y		3505	GAGCGTCGAGTGATCATCAATGAAACGCGCTGCTACACT-CAGTCTGGGTGACAGTGCAGAAG	3563
D	b		5797	GAGCGTCGAGTGAGCCATGATCGCACCACTGCACTCCAGTCTGGTGGTGCAGTAGTACGACT	5856
Q	y		3564	CTGTCTCAAATAATAATAA	3581
D	b		5857	CTGTCTCCAAAAGACAAA	5874
RESULT 13					
AAD52898					
AAD52898 standard; DNA; 47999 BP.					
X	X		AAD52898;		
X	X		14-MAY-2003 (first entry)		
D	E		Human tweety homologue 2 (TTYH2) gene.		
X	X		Human; tweety homologue 2; TTYH2; therapy; cancer; tumour; cytostatic;		
K	W		diagnostic marker; gene; ds.		
X	X		Homo sapiens.		
X	X				
P	H		Key	Location/Qualifiers	
F	T		exon	1936..2074	
F	T			/tag= b	
F	T			/number= 1	
F	T		CDS	1946..45753	
F	T			/tag= a	
F	T			/product= "Human TTYH2 protein"	
F	T		intron	2075..10376	
F	T			/tag= c	
F	T		exon	10377..10549	
F	T			/tag= d	
F	T			/number= 2	
F	T		intron	10550..16622	
F	T			/tag= e	
F	T		exon	16623..16734	
F	T			/tag= f	
F	T			/number= 3	
F	T		intron	16735..23223	
F	T			/tag= g	
F	T		exon	23224..23444	
F	T			/tag= h	
F	T			/number= 4	
F	T		intron	23445..28299	
F	T			/tag= i	
F	T		exon	28300..28395	
F	T			/tag= j	
F	T			/number= 5	
F	T		intron	28396..28902	
F	T			/tag= k	
F	T		exon	28903..28975	
F	T			/tag= l	
F	T			/number= 6	
F	T		intron	28976..35372	
F	T			/tag= m	
F	T		exon	35373..35442	
F	T			/tag= n	
F	T			/number= 7	
F	T		intron	35443..35705	

FT	exon		/ *tag= o 35706 . 35761
FT			/ *tag= p
FT			/number= 8
FT	intron		35762 . 36266
FT			/ *tag= q
FT	exon		36267 . 36359
FT			/ *tag= r
FT			/number= 9
FT	intron		36360 . 36591
FT			/ *tag= s
FT	exon		36592 . 36684
FT			/ *tag= t
FT			/number= 10
FT	intron		36685 . 38529
FT			/ *tag= u
FT	exon		38530 . 38672
FT			/ *tag= v
FT			/number= 11
FT	intron		38673 . 39376
FT			/ *tag= w
FT	exon		39377 . 39562
FT			/ *tag= x
FT			/number= 12
FT	intron		39563 . 40050
FT			/ *tag= y
FT	exon		40051 . 40129
FT			/ *tag= z
FT			/number= 13
FT	intron		40130 . 45672
FT			/ *tag= aa
FT	exon		45673 . 47158
FT			/ *tag= ab
FT			/number= 14
XX			
FN			WO200292629-A1.
XX			
XX			21-NOV-2002.
XX			
PD			14-MAY-2002; 2002WO-AU000591.
Pf			
XX			
PR			14-MAY-2001; 2001AU-00004971.
XX			
PA			(UYQU-) UNIV QUEENSLAND TECHNOLOGY.
XX			Clements JA;
PI			
XX			
DR			WPI; 2003-129264/12.
DR			P-PSDB; AAE34613.
XX			
XX			New human tweety homolog 2 polypeptides and polynucleotides, useful for producing an antigen-binding molecule that is immuno-interactive with the polypeptide or as diagnostic markers for cancers.
PT			
XX			
PS			Claim 10; Page 128-156; 176pp; English.
XX			
CC			The invention relates to human tweety homologue 2 (TTYH2) polypeptide and polynucleotide sequence. TTYH2 is useful for producing an antigen-binding molecule that is immuno-interactive with the polypeptide. The agent is useful for manufacturing a medicament for restoring a normal level and/or functional activity of TTYH2 expression in a patient, and for treating or preventing cancer or tumour. TTYH2 sequences may also be used to provide both drug targets and regulators to promote or inhibit one or more CC activities, and to provide diagnostic markers for cancers. The present CC sequence is human TTYH2 gene
XX			
SQ			Sequence 47999 BP; 11178 A; 12727 C; 12874 G; 11209 T; 0 U; 11 Other;
			Query Match 3.3%; Score 185.4; DB 8; Length 47999;
			Best Local Similarity 74.9%; Pred. No. 5.5e-25;
			Matches 245; Conservative 0; Mismatches 81; Indels 1; Gaps 1;
QY			3284 AGATGAGTGTGGTGCTCATGCCTGTAATCCCACACTTTTGGGAAGCGGTGGGAGGA 3343

CC The sequence data for this patent did not form part of the printed
 CC specification but was obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences).

XX SQ Sequence 199878 BP; 58961 A; 40585 C; 39746 G; 60450 T; 0 U; 136 Other;

Query Match 3.3%; Score 184.8; DB 10; Length 199878;
 Best Local Similarity 62.9%; Pred. No. 8.5e-25;
 Matches 302; Conservative 0; Mismatches 177; Indels 1; Gaps 1;

Qy 3143 TTGTGTCGCCATTTAAATTCCTCTTCTAGCTTAGTAACTTTAGGATTTTAAATAACA 3202
 Db 159286 TTGGTTACCTTAAATACCAACAACAACCAACTATAATTTGGAGATATTTTCATTAAAC 159227

Qy 3203 ACTATTGAAATCATGACATACGTTTAAATGATATTTTAAATACGTTAGGCTATAAAC 3262
 Db 159226 TCTGATTTCTGGGCGCTGACTGCTAACTTATGCTATGACTCATTAAATAAACAGCAC 159167

Qy 3263 TTTTAAATTTTTTAAAAAATAGATGATGTTGGTGTCTATGCTGTAAATCCCAACACTT 3322
 Db 159166 TTTTTCATCAGAAAGACGCGGGCTGGTGTGTAGCTCATTTCTATAAACCCAGCATTT 159107

Qy 3323 TGGGAAGCGGGTGGGAGGATAGCTTGGTCCAGCAGTTTGACACAGTCAGCGGCAACA 3382
 Db 159106 TGGGACGCTAAGCGCGGAGGATCACTTGGCGCAGGGGTTTCGACAGCCGCTGGGCAACA 159047

Qy 3383 CAGCAAGACCCCATATCTAAAAAACAACAACAATAAAATTTACCTGGGTATGTTGTGC 3442
 Db 159046 TGGCAAGACCCCTCTCTACAAAAAATTTAAAAAATAATAGCCAGGTGGGTGGTTC 158987

Qy 3443 TCACCTGTAGTCCAAGCTACACAGGAAGCTGAGGAGGATCACTTGAGCCAGGAGG 3502
 Db 158986 ATGTTTGTAGTCCAGCTACTCAGGAGGCTGAGGTGGGAGGATCACTTGAGCCAGGAGT 158927

Qy 3503 TTGAGGCTGAGTGCATCCATCAAGACCGCTGCTACT- CAGTCTGGGTGACAGTGCAGA 3561
 Db 158926 TTAAGGCTGAGTGAATATGATCACCCTCTGCACCTCCAGCCTGGGTGACAGAGTAAGA 158867

Qy 3562 AGCTGTCTCAAAAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 3621
 Db 158866 CCCATTTCTAAGAAAAAAGAATTTGAAGTAAAAATTTTAAATATTAAGAAGGGAAA 158807

RESULT 15
 ADC24703/c
 ID ADC24703 standard; DNA; 335199 BP.
 XX AC ADC24703;
 XX DT 18-DEC-2003 (first entry)
 XX DE Human wild type HNL4Y genomic sequence.
 XX KW ds; nootropic; neuroleptic; tranquilizer; gene therapy; synaptogenesis;
 KW mutation; neurological disease; mental disorder; psychiatric illness;
 KW autism; Asperger syndrome; schizophrenia;
 KW attention deficit hyperactivity disorder.
 XX OS Homo sapiens.
 XX FH Key Location/Qualifiers
 FT exon 10156..10298
 FT /tag= a
 FT /number= 1b
 FT intron 10299..108499
 FT /tag= b
 FT /number= 1b
 FT exon 108500..109001
 FT /tag= c
 FT /number= 2
 FT intron 109002..205868
 FT /tag= d
 FT /number= 2

FT exon 205869..205928
 FT /tag= e
 FT /number= 2b
 FT intron 205929..209525
 FT /tag= f
 FT /number= 2b
 FT exon 209526..209679
 FT /tag= g
 FT /number= 3
 FT intron 209680..235027
 FT /tag= h
 FT /number= 3
 FT exon 235028..235139
 FT /tag= i
 FT /number= 3b
 FT intron 235140..238089
 FT /tag= j
 FT /number= 3b
 FT exon 238090..238212
 FT /tag= k
 FT /number= 3t
 FT intron 238213..310596
 FT /tag= l
 FT /number= 3t
 FT exon 310597..310783
 FT /tag= m
 FT /number= 4
 FT intron 310784..316138
 FT /tag= n
 FT /number= 4
 FT exon 316139..316929
 FT /tag= o
 FT /number= 5
 FT intron 316930..326821
 FT /tag= p
 FT /number= 5
 FT exon 326822..330136
 FT /tag= q
 FT /number= 6
 PN W02003045998-A2.
 XX 05-JUN-2003.
 XX 02-DEC-2002; 2002WO-FR004134.
 XX 30-NOV-2001; 2001CA-02364106.
 XX (INRM) INSERM INST NAT SANTE & RECH MEDICALE.
 XX (INSP) INST PASTEUR.
 XX (ASSI-) ASSISTANCE PUBLIQUE HOPITAUX PARIS.
 XX Bourgeron T, Jamain S, Quach H, Betancur C, Leboyer M;
 PI Gillberg C;
 XX WPI; 2003-493399/46.
 XX New nucleic acid encoding mutant protein involved in synaptogenesis, and
 FT useful for treatment and diagnosis of e.g. autism, Asperger syndrome, and
 FT schizophrenia.
 XX Claim 10; SEQ ID NO 4; 416pp; French.
 XX The invention relates to an isolated or purified polynucleotide encoding
 CC a polypeptide (the wild-type form of which is involved in synaptogenesis)
 CC that includes at least one mutation associated with development of
 CC neurological disease and/or a predisposition to development of mental
 CC disorders or psychiatric illness. The polypeptide are used to screen for
 CC agents that modulate their activity. Also nucleic acid, polypeptide,
 CC polypeptide-specific antibodies, vectors containing he nucleic acid and
 CC host cells containing the vector, are useful as pharmaceuticals for
 CC treating mental and neurological disorders, specifically autism, Asperger
 CC syndrome, schizophrenia and attention deficit hyperactivity disorder. The

CC wild-type forms of the nucleic acid and polypeptide can be used
CC similarly. Also detecting mutations in the nucleic acid and polypeptide,
CC or measuring activity of the polypeptide, can be used to detect
CC biochemical disorders that affect formation of synapses and to diagnose
CC mental disease. This sequence corresponds to the genomic sequence of the
CC human wild type HNL4Y gene.
XX

SQ Sequence 335199 BP; 98958 A; 63589 C; 65677 G; 106975 T; 0 U; 0 Other;

Query Match 3.2%; Score 183; DB 10; Length 335199;

Best Local Similarity 65.2%; Pred. No. 2e-24;

Matches 303; Conservative 0; Mismatches 155; Indels 7; Gaps 2;

Qy	3158	TTTCTGTTCCCTCTAGCTTAGTAACTTTAGGATTTTATATACAACTATTGAAATCATG	3217
Db	239974	TTTATATATCTTATATAAAGTCATCTTTTATATATATAAAATATATTATATATATAATT	239915
Qy	3218	ACATACGTTTAAATGATATTATTAAATACGTTAGGCTATATAAACCTTTTAAATTTTTTAA	3277
Db	239914	ATGTTACAAAATCTAATATATATATAATAATAAATCATTCAGTGACATTGAAGAGAAAA	239855
Qy	3278	AAAAATAGATGAGTGGTGGCTCATGCCCTGTAAATCCCAACACCTTTGGGAAGCGGGTCG	3337
Db	239854	AAAAATGGCCAGGTGCGGTGCTCATGCCCTGTAAATCCCAACACCTTTGGGAGGCGGAGTG	239795
Qy	3338	GGAGGATAGCTTGAGTCCAGCAGTTTGGAGACCAGTCAGGGCCACACAGCAAGACCCCAT	3397
Db	239794	GGTGGATTGCTGAGCTCAGGAGTTTGAGACCAGCCCTGGGCAACACAGTAAACCCCATC	239735
Qy	3398	TCTAAAAAAACAAAAACAAAAATTTACCTGGGTATGGTTGTGCTCACCTGTAGTCCAA	3457
Db	239734	TCTATGGAATACAA-----AAAAATTAGTGGCGTGGTGGTGCACACCTGTAGTCCCA	239681
Qy	3458	GCTACACAGGAAGCTGAGGAGAGGATCACTTGAGCCCAAGAGGTTGAGGCTGCAGTGA	3517
Db	239680	GCTACTGGGAGGCTGAGACAGGAGAATTGCTTGAACCCAGAGGCGAGAGGTTGCAGTGA	239621
Qy	3518	TCCATGAACGGCTGTACACT-CAGTCTGGGTGACAGTGCAGNAGCTGTCTCAAAAT	3576
Db	239620	GCTGAGATCAGCCATTGCACTCCAGCTGGGTGACAGAGAGACTCTGTCTCAAAAAA	239561
Qy	3577	AATAAATAAATAAATAAATAAATTTTAAAAAACAATAAATAAATAAATAAATAAATAA	3621
Db	239560	ATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA	239516

Search completed: June 23, 2005, 01:20:03
Job time : 2733 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 23, 2005, 00:25:12 ; Search time 845 Seconds
(without alignments)
10991.132 Million cell updates/sec

Title: US-10-030-294-1
Perfect score: 5676
Sequence: 1 cagctgttcaggatgctg.....tttataaattgtctccgt 5676

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA: *
1: /cgn2_6/ptodata/1/ina/5A COMB.seq: *
2: /cgn2_6/ptodata/1/ina/5B COMB.seq: *
3: /cgn2_6/ptodata/1/ina/6A COMB.seq: *
4: /cgn2_6/ptodata/1/ina/6B COMB.seq: *
5: /cgn2_6/ptodata/1/ina/PTUS COMB.seq: *
6: /cgn2_6/ptodata/1/ina/backfiles1.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1987.4	35.0	255679	4	US-09-949-016-17189 Sequence 17189, A
2	190.6	3.4	265038	4	US-09-949-016-15779 Sequence 15779, A
3	187.6	3.3	8133	3	US-09-659-791A-10 Sequence 10, Appl
4	186.4	3.3	601	4	US-09-949-016-142852 Sequence 142852, A
5	186	3.3	17348	4	US-09-949-016-17403 Sequence 17403, A
6	185.6	3.3	601	4	US-09-949-016-201827 Sequence 201827, A
7	185.4	3.3	37292	4	US-09-949-016-15382 Sequence 15382, A
8	184.8	3.3	172677	4	US-09-949-016-13444 Sequence 13444, A
9	184.6	3.3	33272	4	US-09-949-016-16949 Sequence 16949, A
10	184.4	3.2	601	4	US-09-949-016-58699 Sequence 58699, A
11	184	3.2	601	4	US-09-949-016-132156 Sequence 132156, A
12	183.4	3.2	601	4	US-09-949-016-119415 Sequence 119415, A
13	182.8	3.2	36228	4	US-09-949-016-12256 Sequence 12256, A
14	182.8	3.2	36228	4	US-09-949-016-15468 Sequence 15468, A
15	182.6	3.2	134140	4	US-09-949-016-12672 Sequence 12672, A
16	182.6	3.2	134241	4	US-09-949-016-12424 Sequence 12424, A
17	182.6	3.2	134242	4	US-09-949-016-15813 Sequence 15813, A
18	182.6	3.2	134242	4	US-09-949-016-15814 Sequence 15814, A
19	182.6	3.2	134242	4	US-09-949-016-15815 Sequence 15815, A
20	182.4	3.2	601	4	US-09-949-016-132157 Sequence 132157, A
21	182.2	3.2	36651	3	US-09-738-894A-3 Sequence 3, Appli
22	182.2	3.2	36651	4	US-09-964-469-3 Sequence 3, Appli
23	182.2	3.2	51403	4	US-09-949-016-15057 Sequence 15057, A
24	181.8	3.2	4732	4	US-09-949-016-14962 Sequence 14962, A
25	181	3.2	601	4	US-09-949-016-127367 Sequence 127367, A
26	181	3.2	69924	4	US-09-949-016-15367 Sequence 15367, A
27	180.8	3.2	21920	4	US-09-949-016-15609 Sequence 15609, A

c	28	180.2	3.2	86273	4	US-09-949-016-15273	Sequence 15273, A
	29	180.2	3.2	87205	4	US-09-949-016-13430	Sequence 13430, A
	30	180	3.2	78720	4	US-09-949-016-12710	Sequence 12710, A
	31	180	3.2	78720	4	US-09-949-016-17283	Sequence 17283, A
	32	179.6	3.2	103377	4	US-09-949-016-14089	Sequence 14089, A
	33	179.6	3.2	119032	4	US-09-949-016-12160	Sequence 12160, A
	34	179.6	3.2	119032	4	US-09-949-016-17268	Sequence 17268, A
	35	179.6	3.2	163181	4	US-09-949-016-13730	Sequence 13730, A
	36	179.4	3.2	601	4	US-09-949-016-58698	Sequence 58698, A
	37	179.2	3.2	77851	4	US-09-949-016-12508	Sequence 12508, A
	38	179.2	3.2	77867	4	US-09-949-016-13211	Sequence 13211, A
	39	179.2	3.2	77867	4	US-09-949-016-13212	Sequence 13212, A
	40	179.2	3.2	77940	4	US-09-949-016-12509	Sequence 12509, A
	41	179	3.2	65848	4	US-09-949-016-13285	Sequence 13285, A
	42	178.8	3.2	152132	4	US-09-949-016-13845	Sequence 13845, A
	43	178.8	3.2	152145	4	US-09-949-016-12371	Sequence 12371, A
	44	178.6	3.1	186734	4	US-09-949-016-14870	Sequence 14870, A
	45	178.6	3.1	193689	4	US-09-949-016-12350	Sequence 12350, A

ALIGNMENTS

RESULT 1
US-09-949-016-17189
; Sequence 17189, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17189
; LENGTH: 255679
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(255679)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-17189

Query Match	35.0%	Score	1987.4	DB	4	Length	255679
Best Local Similarity	99.9%	Pred. No.	0				
Matches	1999	Conservative	0	Mismatches	1	Indels	1
							Gaps
Qy	3676	TATCCCTGGAGTAAATTTAAATATTTAGAGAGTTCTTCTCATTTCTTAGAGAGCTC	3735				
		1 TATCCCTGGAGTAAATTTAAATATTTAGAGAGTTCTTCTCATTTCTTAGAGAGCTC	60				
Db							
Qy	3736	GAATTTGTAATATATCAGAGCTAGAGGAACTAGGGCTCGCCACTCCAAAGTGTGGTCCA	3795				
		61 GAATTTGTAATATATCAGAGCTAGAGGAACTAGGGCTCGCCACTCCAAAGTGTGGTCCA	120				
Db							
Qy	3796	AGGACCAAGAGCATCAAGTAACTGGGAACTGTAGAAATGAGAGTCTTAGGCCTCAC	3855				
		121 AGGACCAAGAGCATCAAGTAACTGGGAACTGTAGAAATGAGAGTCTTAGGCCTCAC	180				
Db							
Qy	3856	CCGACGACTACTGAACCAAGATCTGCATTACAGATTTCTAGGTGCTCTACGGGCACAT	3915				
		181 CCCGACGACTACTGAACCAAGATCTGCATTACAGATTTCTAGGTGCTCTACGGGCACAT	240				
Db							
Qy	3916	TAAACCTTGAGAGCTCTGCAGTAAATCTTCACTCCACTTTTCATTATAATGAATC	3975				

Db 241 TAAAACTGAGAACTCTGCACTAGAAATCTTCACTCCACCTTTCATTATAAATGGAATC 300
QY 3976 ACTTGGGCTGTGGTACAGGAATATGATTATTTTAAATTCAGAACTTCTATTAGGTC 4035
Db 301 ACTT-GGGCTGTGGTACAGGAATATGATTATTTTAAATTCAGAACTTCTATTAGGTC 359
QY 4036 ATCTATATTTGCTAATAGCAGGGAAGAAAGCCAAATCTTTTAACTGCAATTAACAATCT 4095
Db 360 ATCTATATTTGCTAATAGCAGGGAAGAAAGCCAAATCTTTTAACTGCAATTAACAATCT 419
QY 4096 ATAATTAATTAAGTAAACAATCTTCCCTTTAAGTTTACATTTTGTGGACAAGCTGTTT 4155
Db 420 ATAATTAATTAAGTAAACAATCTTCCCTTTAAGTTTACATTTTGTGGACAAGCTGTTT 479
QY 4156 GATTGGCTGGGCTCAGGCGGGCTGTTTGTGAATTTCAAAATTCACAGATGTTAGCCG 4215
Db 480 GATTGGCTGGGCTCAGGCGGGCTGTTTGTGAATTTCAAAATTCACAGATGTTAGCCG 539
QY 4216 CTCTCGGGCTAAGTAAAGGAAGAAATGTCAAGTTTAAATAGCTTCCCTTCCATCCT 4275
Db 540 CTCTCGGGCTAAGTAAAGGAAGAAATGTCAAGTTTAAATAGCTTCCCTTCCATCCT 599
QY 4276 GGCTGAAGCAACAATAAATATTTTATGAACAACATTTTGTAGTATTTACTTACAG 4335
Db 600 GGCTGAAGCAACAATAAATATTTTATGAACAACATTTTGTAGTATTTACTTACAG 659
QY 4336 GGAATGTCAAAATTTCTCTGAAGGGCTTTAGATTGTCTCAAACTTGCATCTACTGA 4395
Db 660 GGAATGTCAAAATTTCTCTGAAGGGCTTTAGATTGTCTCAAACTTGCATCTACTGA 719
QY 4396 TGTCAACCTATTACAGTGTGTCTGTGACTAGGGGTGAAGGAAGATGTGAACCTACC 4455
Db 720 TGTCAACCTATTACAGTGTGTCTGTGACTAGGGGTGAAGGAAGATGTGAACCTACC 779
QY 4456 ATGTTAGTACCGTTAGATACACAGAGTGGTTTTTTTCCCTGTTGGAGTCTATCCTA 4515
Db 780 ATGTTAGTACCGTTAGATACACAGAGTGGTTTTTTTCCCTGTTGGAGTCTATCCTA 839
QY 4516 ACTGAGCTTCTGAATCATATTTTCATTTCAAAATTCACAAACAGGATAGATTTA 4575
Db 840 ACTGAGCTTCTGAATCATATTTTCATTTCAAAATTCACAAACAGGATAGATTTA 899
QY 4576 CAGCCCATATTCAGAAAGGAATAAATTTATTTGTGTAGTACTTTCCTCATATTACACT 4635
Db 900 CAGCCCATATTCAGAAAGGAATAAATTTATTTGTGTAGTACTTTCCTCATATTACACT 959
QY 4636 GATTTGGGAATATATGAACAATTTTATGGTTTCTTTTCGAAGTAGGTCAAGTCAAAAGCAA 4695
Db 960 GATTTGGGAATATATGAACAATTTTATGGTTTCTTTTCGAAGTAGGTCAAGTCAAGCAA 1019
QY 4696 AACCAAAAACAGCAAAAACCTGTAGACATTAAGATAGATAGAGGCGGACTGAGAGATTA 4755
Db 1020 AACCAAAAACAGCAAAAACCTGTAGACATTAAGATAGATAGAGGCGGACTGAGAGATTA 1079
QY 4756 AATAAATCTAGATATTTTATTAACAGGCAATTTGAAATTAATTTGTCACITTCAGATA 4815
Db 1080 AATAAATCTAGATATTTTATTAACAGGCAATTTGAAATTAATTTGTCACITTCAGATA 1139
QY 4816 TTCTCAATATATATATTTTCAATTTTAAATATCTTTAAGAAAATTTACTATATTATATG 4875
Db 1140 TTCTCAATATATATATTTTCAATTTTAAATATCTTTAAGAAAATTTACTATATTATATG 1199
QY 4876 TAAGTACATGTGATGTTTGTAGGTAGGATATTTTAACTCAATAAAGGTTATTTCTTTT 4935
Db 1200 TAAGTACATGTGATGTTTGTAGGTAGGATATTTTAACTCAATAAAGGTTATTTCTTTT 1259
QY 4936 ATTGGGTGAGGCAAGCTTCTAAGGGGATGTGAAGGGATATCTCTTCTCTAGCTGA 4995
Db 1260 ATTGGGTGAGGCAAGCTTCTAAGGGGATGTGAAGGGGATATCTCTTCTCTAGCTGA 1319
QY 4996 GAGGAAGAGTGAGTTCTTAAGTTAAATATTAATCAAGGAATTTCCCTGTCTTGTCTATTGA 5055

Db 1320 GAGGAAGAGTGAGTTCTTAAGTTAAATATTAATCAAGGAATTTCCCTGTCTTTGTCTATTGA 1379
QY 5056 GATTGTGACACAAACAGGCGGTTGGCTGAAAGGAACTTGAAGGCGGGGAGGAGGAA 5115
Db 1380 GATTGTGACACAAACAGGCGGTTGGCTGAAAGGAACTTGAAGGCGGGGAGGAGGAA 1439
QY 5116 ATAGATGAAAAAACAACAAACAAACAAACCTTCCCTAAGCAGCTCTACAAACATTTTATGC 5175
Db 1440 ATAGATGAAAAAACAACAAACAAACAAACCTTCCCTAAGCAGCTCTACAAACATTTTATGC 1499
QY 5176 CCAGAAAATAGTACAGAAAATCTCTCAAAATCAAAACCAAGTATCCAGATCAAGGAAGTTTA 5235
Db 1500 CCAGAAAATAGTACAGAAAATCTCTCAAAATCAAAACCAAGTATCCAGATCAAGGAAGTTTA 1559
QY 5236 TGTAGCTGAGCAGGAGGAGGAGCAGCTCATCAGCTCAGTTTCAAAAGTCCAGGCTGCT 5295
Db 1560 TGTAGCTGAGCAGGAGGAGGAGCAGCTCATCAGCTCAGTTTCAAAAGTCCAGGCTGCT 1619
QY 5296 GAAATTTAAATCTCTGATGCCATTTTCATGCCAGCATCCAAATCACGACAGATCAGAACTTCA 5355
Db 1620 GAAATTTAAATCTCTGATGCCATTTTCATGCCAGCATCCAAATCACGACAGATCAGAACTTCA 1679
QY 5356 GAGATCCCTCCAGCTCCAAATTTGCCAAACAAAGTGTGGCTACTATACGTCAAGGACTCT 5415
Db 1680 GAGATCCCTCCAGCTCCAAATTTGCCAAACAAAGTGTGGCTACTATACGTCAAGGACTCT 1739
QY 5416 GAAGCGCTGAGAGGAGGGAAGAAACAAAGTGTGGCTACTATACGTCAAGGACTCT 5475
Db 1740 GAAGCGCTGAGAGGAGGGAAGAAACAAAGTGTGGCTACTATACGTCAAGGACTCT 1799
QY 5476 GTGTTTATCAAGTTTGTAGTCAATTTGATGAATCTCATTTGGCTAAAAATCAAGAAACGCTCCG 5535
Db 1800 GTGTTTATCAAGTTTGTAGTCAATTTGATGAATCTCATTTGGCTAAAAATCAAGAAACGCTCCG 1859
QY 5536 CCTTTTGCAGATATGTATGAAGGAGAGAGTGCCTTAAACTTCTATGTCTGATAGCATTT 5595
Db 1860 CCTTTTGCAGATATGTATGAAGGAGAGAGTGCCTTAAACTTCTATGTCTGATAGCATTT 1919
QY 5596 GACCTATTGCTTTTGTAGCTCCGCGCTTATATCTATATATACACAGGTTATTTGTCTATA 5655
Db 1920 GACCTATTGCTTTTGTAGCTCCGCGCTTATATCTATATATACACAGGTTATTTGTCTATA 1979
QY 5656 TTTTATATAATTTGTTCTCCGT 5676
Db 1980 TTTTATATAATTTGTTCTCCGT 2000

RESULT 2

US-09-949-016-15779
; Sequence 15779, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15779
; LENGTH: 265038
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(265038)


```

US/09-949-016-17403
Sequence 17403, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN
SEQUENCES OF HUMAN DNA WITH HUMAN DISEASE,
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20

```



```
; SEQ ID NO 58699
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-58699

Query Match
  3.2%; Score 184.4; DB 4; Length 601;
Best Local Similarity 69.3%; Pred. No. 6.1e-33;
Matches 280; Conservative 1; Mismatches 117; Indels 6; Gaps 2;

QY 3195 AAATAACAACCTATTGAATCATGACATACGTTAAATGATATATTAAATACGTTAGGC 3254
Db 156 AATTTAAAAATATACAAATCATACAAATAGATGTTGAGTTAAATTCAAATTCAGTTGAC 215
QY 3255 TATAAACCTTTTAAATTTTAAAAAATAGAT-----GAGTGTGGTGGCTCATGCTGT 3309
Db 216 TTTTGTGATTTGTTATTTTAAAAAAGTTGACCATGTGTGGTGGCTGTGTGT 275
QY 3310 AATCCCAACACTTTTGGGAAGCGGGTGGGAGTAGCTTGAGTCCAGCAGTTTGAGACC 3369
Db 276 AATCCAGCACTTTGGAAAGCTGAGYTGGAAGTAGCTTGAGACCAGGAGTTGAGACC 335
QY 3370 AGTCAGGCAACACAGCAAGACCCATATCTAAAAAACAACAAACAAAAATTACCTG 3429
Db 336 AGCCTGGCAACATGGCAAGACCTCATCTCTACTAAAAATAAAAAATAAAGCTA 395
QY 3430 GGTATGTTGTGCTCACTGTAGTCCAGCTTACACAGGAAGCTGAGGCAGAGGATCACT 3489
Db 396 GGCATAGCAGTGTGTTCATATAGTCCAGCTACTCCAGGAGCAGAGTGGGAGGATCGCT 455
QY 3490 TGAGCCAGGAGGTTGAGGCTGCACTGATCCATGAACGCGTGTCACT-CAGTCTGGG 3548
Db 456 TGAACCCAGAGGTTGAGGCTGCACTGATGAGTGTGATCACCACTGCACTCCAGCCTAGG 515
QY 3549 TGACAGTGCAGAGGCTGTCTCAAAAAATAAATAAATAAAAA 3592
Db 516 TGACAGAGAAAGACCTTGTCTCAAGAAATAAGCAACAAAAATA 559

RESULT 11
US-09-949-016-132156
; Sequence 132156, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR FILING DATE: 2000-10-03
; PRIOR FILING DATE: 2000-10-03
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 132156
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-132156

Query Match
  3.2%; Score 184; DB 4; Length 601;
Best Local Similarity 62.4%; Pred. No. 7.5e-33;
Matches 322; Conservative 1; Mismatches 186; Indels 7; Gaps 2;

QY 3195 AAATAACAACCTATTGAATCATGACATACGTTAAATGATATATTAAATACGTTAGGC 3254
Db 5 ACACAACAACAACAAAAAGATAAAGTGAAGCTTTTATTAAGGTTAAAAAAGTTGAACCTCAA 64
QY 3255 TATAAACCTTTTAAATTTTAAAAAATAGATGAGTGTGGTGCCTCATGCTGTAATCC 3314

; SEQ ID NO 58999
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-58999

Query Match
  3.2%; Score 183.4; DB 4; Length 601;
Best Local Similarity 62.2%; Pred. No. 1e-32;
Matches 304; Conservative 1; Mismatches 182; Indels 2; Gaps 1;

QY 3143 TTGTGGCCCAATTAATTTCTGTCTCTTAGCTTAGTAAGTATTTAGGATTTTAAATAACA 3202
Db 47 TTTTCTTTCTTTTGTCTTATTTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 106
QY 3203 ACTATTGAATCATGACATACGTTTAAATGATATATTAAATACGTTAGGCTTATAAACC 3262
Db 107 CTTTATTAATTAAGAAAAAATTTTAAAGTTTATTTTCAGGACACCATTTGAATTAATAC 166
QY 3263 TTTTAAATTTTAAAAAATAGATGAGTGTGGTGCCTCATGCTGTATATCCCAACACTT 3322
Db 167 TGAATAAATAAATAAATAAGAGGCTGGGTATAGTGGCTTATGCTTATATATATATATATC 226

; SEQ ID NO 58999
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-58999

Query Match
  3.2%; Score 183.4; DB 4; Length 601;
Best Local Similarity 62.2%; Pred. No. 1e-32;
Matches 304; Conservative 1; Mismatches 182; Indels 2; Gaps 1;

QY 3143 TTGTGGCCCAATTAATTTCTGTCTCTTAGCTTAGTAAGTATTTAGGATTTTAAATAACA 3202
Db 47 TTTTCTTTCTTTTGTCTTATTTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 106
QY 3203 ACTATTGAATCATGACATACGTTTAAATGATATATTAAATACGTTAGGCTTATAAACC 3262
Db 107 CTTTATTAATTAAGAAAAAATTTTAAAGTTTATTTTCAGGACACCATTTGAATTAATAC 166
QY 3263 TTTTAAATTTTAAAAAATAGATGAGTGTGGTGCCTCATGCTGTATATCCCAACACTT 3322
Db 167 TGAATAAATAAATAAATAAGAGGCTGGGTATAGTGGCTTATGCTTATATATATATATATC 226
```

3323	Qy	TGGGAAGCCGGGTCGGAGGATAGCTTGAGTCAGACAGTTTGAGACACGATCAGGCGACACA	3382
227	Db	TGGGAGCCGAGGTGGGTGGATCACATTGAGGCCAGGAGTTTCAGAGACACAGCCTGGGTGACA	286
3383	Qy	CAGCAAGACCCCATATCTATAAAAAACAAAAACAAAACAAAATTACCTGGGTATGTTGTGC	3442
287	Db	TAGTGAGATCTTGTTMTCTGCAAAAAAATAAATGTTTAAAAATCAGCCAGGTGTGGTGGTGC	346
3443	Qy	TCACTCTGTAGTCCAAGCTACACAGGAAGCTGAGCGAGAAGGATCATTGAGGCCACGAGG	3502
347	Db	CTCCCTGTAGTTTCAGATCTCGAGAGGCTTAAGGTGGAGGATGCTTTGAGTCCAGAGT	406
3503	Qy	TTGAGGCTGCAGTGCATCCATGAACGGCTCTCTACACT--CAGTCTGGGTGACAGTGCAG	3560
407	Db	TTGAGGCTGCAGTAAGCTATGATCATGCCATTGCAGTTTCAGCTTTGGGTACAGAGCAAG	466
3561	Qy	AAGCTGTCTCAAAAAATAATAATAATAAAAAATACTTTTAAAAAACAATAATAA	3620
467	Db	ACCTGTCTCAACAAAAACACAGAAGTATGTAATAATGATATATAGTCCAGGTATTATTTA	526
3621	Qy	ATTTTAAAA	3629
527	Db	ACAGTAATA	535

```

RESULT 13
US-09-949-016-12256/c
; Sequence 12256, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C1001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12256
; LENGTH: 36228
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-12256

```

[illegible]

Qy	3495	CCAGGAGGTTGAGGCTCCAGTGCATGAACGGCGTCTACACT-CAGTTCGGGTGACA	3553
Db	6882	CCGGGAGGCAGAGGTTGCAGTGAACCGGCTTGCACCTCAGCGCAGGCGGACA	6823
Qy	3554	GTCCMAGAACGCTGCTCAGAAAAATAATAATAAAAAATAAATTTTAAAAAAACAAAAAT	3613
Db	6822	GAGCAAAATCTCCATCTCTCAAAAAAATAAAAAAATAAAAAAATAAAAAATGAAAAACACAC	6763
Qy	3614	TAATTAATTTTAAAAACACAAACACACACTAGAGATGTTTGCAAAATTTGATTTATTTGGGAGTC	3673
Db	6762	TCAATGAATAGGATAAAATATTTTGCAAATCATATATCTGATAAGGACCTGTATCTAGAA	6703
Qy	3674	TATATCCCTCGAAGTTAATTTAAAAATATTTAGAAAGA	3709
Db	6702	TATAAAAAGAAATCTCTGTAAGTCAATTAATAAGAAAGA	6667

RESULT 14
US-09-949-016-15468/c
Sequence 15468, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIORITY APPLICATION NUMBER: 60/241,755
PRIORITY FILING DATE: 2000-10-20
PRIORITY APPLICATION NUMBER: 60/237,768
PRIORITY FILING DATE: 2000-10-03
PRIORITY APPLICATION NUMBER: 60/231,498
PRIORITY FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 15468
LENGTH: 36228
TYPE: DNA
ORGANISM: Human
US-09-949-016-15468

	Query Match	3.2%	Score 182.8;	DB 4;	Length 36228;
	Best Local Similarity	62.4%;	Pred. No. 7.9e-32;		
	Matches 322;	Conservative 0;	Mismatches 187;	Indels 7;	Gaps 2;
Qy	3195	AAATACAACTATTGAAATCATGACATACGTTTAAATGATATTATTAAATACGTTAGGC	3254		
Db	7176	ACACAAACAACAAAAAGATAAATCTGGACTTTATTAAAGTTAAAAACCTTGAACTCAA	7117		
Qy	3255	TATTAACCTTTTAAATTTTTTAAAAAATAGATGAGTGTGGTGGCTCATGCGCTGTAATCC	3314		
Db	7116	AAGTACCATAGAAAAATGAAAAAGCGAGCGCGAGTGCAGTGGCTCACGCTGTAATCC	7057		
Qy	3315	CAACACTTTGGGAAGCGGGTCGGAGGATAGCTTGAGTCCAGAGTTTGAGACCCAGTCA	3374		
Db	7056	CACCACTTTGGGAGCGGAGCGGGGGATTGCTTGAGGCGAGAGTTCCAGAGCCAGTCT	6997		
Qy	3375	GGGCAACACAGCAAGACCCCATATCTTAAAAAACAACAAACAAATTAACCTGGGTAT	3434		
Db	6996	GGCCAAATGGCAAAACCCCTGTCTCT-----ACTAAAAATACAAAATTAGCCTGTGT	6943		
Qy	3435	GGTTGTGCTCACCTGTAGTCCAAAGCTACACAGAAAGCTGAGGCAGAGGATCACTTGAGC	3494		
Db	6942	GGTGGCGGACCTGTATNTCCAGTTACTCAGAGGCTGAGGCAGAGNATCGCTGAAC	6883		
Qy	3495	CCAGGAGGTTTGAGCTGCACTGATCCATGAACCGCTGTCACT-CAGTCTGGGTGACA	3553		
Db	6882	CCGGAGGCGAGAGGTTGACGTGAGCCAAAGATCACGCCGTTTGCACTCCAGCCAGGGCGACA	6923		
Qy	3554	GTGCAAGAAGCTGTCTCAAAAATAAATAAATAAATAAATAAATTAACCTTTTAAAAACAAAAAT	3613		

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 23, 2005, 00:34:38 ; Search time 3068 Seconds
(without alignments)
11563.865 Million cell updates/sec

Title: US-10-030-294-1

Perfect score: 5676

Sequence: 1 cagctgttcaggatgctg.....tttataaattgttctccgt 5676

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 6067389 seqs, 3125258755 residues

Total number of hits satisfying chosen parameters: 12134778

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
- 14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
- 17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
- 18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:*
- 19: /cgn2_6/ptodata/1/pubpna/US10G_PUBCOMB.seq:*
- 20: /cgn2_6/ptodata/1/pubpna/US10H_PUBCOMB.seq:*
- 21: /cgn2_6/ptodata/1/pubpna/US10I_PUBCOMB.seq:*
- 22: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
- 23: /cgn2_6/ptodata/1/pubpna/US11A_PUBCOMB.seq:*
- 24: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*
- 25: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
- 26: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	196.2	3.5	65793	21 US-10-703-817-3	Sequence 3, Appli
C 2	196	3.5	304905	18 US-10-271-416-1	Sequence 1, Appli
C 3	193	3.4	13070	21 US-10-741-600-17605	Sequence 17605, A
C 4	193	3.4	156318	21 US-10-741-600-17574	Sequence 17574, A
C 5	188.4	3.3	135005	20 US-10-723-860-2320	Sequence 2320, Ap
C 6	187.6	3.3	8133	18 US-10-380-124-10	Sequence 10, Appl
C 7	186	3.3	18861	11 US-09-984-429-513	Sequence 513, App

C 8	185.6	3.3	611	13 US-10-027-632-181221	Sequence 181221,
C 9	185.6	3.3	611	17 US-10-027-632-181221	Sequence 181221,
C 10	185.4	3.3	3287	13 US-10-027-632-115440	Sequence 115440,
C 11	185.4	3.3	3287	17 US-10-027-632-115440	Sequence 115440,
C 12	183	3.2	335199	21 US-10-496-011-4	Sequence 4, Appli
C 13	182.6	3.2	47903	19 US-10-322-281-808	Sequence 808, App
C 14	182.4	3.2	736	13 US-10-027-632-17813	Sequence 17813, A
C 15	182.4	3.2	736	17 US-10-027-632-17813	Sequence 17813, A
C 16	182.2	3.2	36651	9 US-09-964-469-3	Sequence 3, Appli
C 17	182.2	3.2	36651	16 US-10-425-962-3	Sequence 3, Appli
C 18	182	3.2	55735	21 US-10-741-600-17698	Sequence 8, Appli
C 19	181.8	3.2	50000	17 US-10-364-505-8	Sequence 8, Appli
C 20	181.8	3.2	50000	19 US-10-681-199-8	Sequence 8, Appli
C 21	181.4	3.2	145806	20 US-10-719-993-6943	Sequence 6943, Ap
C 22	181	3.2	58985	10 US-09-901-152-3	Sequence 3, Appli
C 23	181	3.2	58985	21 US-10-483-593-3	Sequence 3, Appli
C 24	181	3.2	143601	10 US-09-855-824-3	Sequence 3, Appli
C 25	181	3.2	143601	20 US-10-476-542-3	Sequence 110324,
C 26	180.8	3.2	2214	13 US-10-027-632-110324	Sequence 110324,
C 27	180.8	3.2	2214	17 US-10-027-632-110324	Sequence 110324,
C 28	180.8	3.2	198522	13 US-10-087-192-244	Sequence 244, App
C 29	180.6	3.2	13487	20 US-10-719-993-6919	Sequence 6919, Ap
C 30	180.2	3.2	95832	21 US-10-741-600-17662	Sequence 17662, A
C 31	180	3.2	94720	18 US-10-052-482-160	Sequence 160, App
C 32	179.8	3.2	108316	17 US-10-292-798-1789	Sequence 1789, Ap
C 33	179.8	3.2	108317	15 US-10-017-161-2143	Sequence 2143, Ap
C 34	179.6	3.2	51657	17 US-10-057-4758-10475	Sequence 10475, A
C 35	179.6	3.2	51657	17 US-10-154-884B-10475	Sequence 10475, A
C 36	179.6	3.2	65608	9 US-09-962-436-292	Sequence 292, App
C 37	179.6	3.2	65608	9 US-09-962-832-119	Sequence 119, App
C 38	179.6	3.2	65608	9 US-09-954-531-180	Sequence 180, App
C 39	179.6	3.2	65608	21 US-10-843-641A-1247	Sequence 1247, Ap
C 40	179.6	3.2	65608	21 US-10-843-641A-2751	Sequence 2751, Ap
C 41	179.6	3.2	65608	21 US-10-843-641A-6005	Sequence 6005, Ap
C 42	179.4	3.2	68744	20 US-10-719-993-6752	Sequence 6752, Ap
C 43	179.4	3.2	160556	20 US-10-719-993-6827	Sequence 6827, Ap
C 44	179.2	3.2	812	13 US-10-027-632-170028	Sequence 170028,
C 45	179.2	3.2	812	17 US-10-027-632-170028	Sequence 170028,

ALIGNMENTS

RESULT 1
US-10-703-817-3/c
; Sequence 3, Application US/10703817
; Publication No. US20050118117A1
; GENERAL INFORMATION:
; APPLICANT: ROTH, RICHARD B.
; APPLICANT: NELSON, MATTHEW ROBERTS
; APPLICANT: KAMMERER, STEFAN M.
; APPLICANT: BRAUN, ANDREAS
; TITLE OF INVENTION: METHODS FOR IDENTIFYING RISK OF MELANOMA AND TREATMENTS
; FILE REFERENCE: SEQ-4061-UT
; CURRENT APPLICATION NUMBER: US/10703.817
; CURRENT FILING DATE: 2003-11-06
; PRIOR APPLICATION NUMBER: 60/489,703
; PRIOR FILING DATE: 2003-07-23
; PRIOR APPLICATION NUMBER: 60/424,475
; PRIOR FILING DATE: 2002-11-06
; NUMBER OF SEQ ID NOS: 255
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 3
; LENGTH: 65793
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-703-817-3

Query Match 3.5%; Score 196.2; DB 21; Length 65793;
Best Local Similarity 70.0%; Pred. No. 4e-30;
Matches 278; Conservative 0; Mismatches 118; Indels 1; Gaps 1;

; CURRENT FILING DATE: 2003-12-22
 ; NUMBER OF SEQ ID NOS: 73997
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 17574
 ; LENGTH: 156318
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; LOCATION: (1)_(156318)
 ; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-
 US-10-741-600-17574

Query Match 3.4%; Score 193; DB 21; Length 156318;
 Best Local Similarity 70.8%; Pred. No. 2.9e-29;
 Matches 271; Conservative 0; Mismatches 110; Indels 2; Gaps 1;
 QY 3210 AATCATGACATAGCTTTAAATGATATATTTAAATACGTTAGCTATAACCTTTTAAA 3269
 Db 66635 ATATCTGTACAAAACATCTGATTTTCACTTTTCAAAATCTTTAGGTAGTAAAGTTAAATC 66576
 QY 3270 TTTTAAAAAATAGATGATGTTGGCTCATGCTGTAATCCCAACACTTTGGGAAG 3329
 Db 66575 TAAAGGTCAAGATGTTGGGTATGGTATGCTACGCTGTAATCCCAACACTCTAGGAGG 66516
 QY 3330 CCGGGTCGGGAGGATAGCTTGAGTCCAGCAGTTTGGAGCAGCTCAGGGCAACACAGCAAG 3389
 Db 66515 CCGAGGCGGGCATATCGCTTGAGCCGAGGGTTGAGACCAGCTAGGGCAACAGGCAA 66456
 QY 3390 ACCCATATCTAAAAAACAACAAAAATACCTGGTATGTTGTGCTACCTG 3449
 Db 66455 ACCCATCTCTACAAAATAAAATACA-ATAATGAGCTGGGCATGGTGGCAGCACCTG 66398
 QY 3450 TAGTCCAGCTTACACAGAGCTGAGCAGAGGATCACTTGACCCAGGAGGTGAGGC 3509
 Db 66397 TGGTCCAGCTACTTGGAGGCTGAGCAGGAGGATGGCTTGAGCCAGGAGGTGGGC 66338
 QY 3510 TGAGTGCATCATCAACGCTGCTACACTCAGTCTGGTGACAGTGCAGAGAGCTGCT 3569
 Db 66337 TGCAGTGCAGTGCAGTGTACTGTACTGTACTGTACTGTACTGTACTGTACTGTACTGT 66278
 QY 3570 CAAAAATAATAATAATAATAA 3592
 Db 66277 CAAAAATAATAATAATAATAA 66255

RESULT 5
 US-10-723-860-2320
 ; Sequence 2320, Application US/10723860
 ; Publication No. US20040253606A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Aziz, Natasha
 ; APPLICANT: Ginsburg, Wendy M.
 ; APPLICANT: Zlotnik, Albert
 ; TITLE OF INVENTION: Methods of Diagnosis of Soft Tissue Sarcoma, Compositions &
 ; TITLE OF INVENTION: Methods for Screening for Soft Tissue Sarcoma Modulators
 ; FILE REFERENCE: 05882.0193.NPUS01
 ; CURRENT APPLICATION NUMBER: US/10/723,860
 ; CURRENT FILING DATE: 2003-11-26
 ; PRIOR APPLICATION NUMBER: 60/429,739
 ; PRIOR FILING DATE: 2002-11-26
 ; NUMBER OF SEQ ID NOS: 8393
 ; SOFTWARE: Patent in version 3.2
 ; SEQ ID NO 2320
 ; LENGTH: 135005
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-10-723-860-2320

Query Match 3.3%; Score 188.4; DB 20; Length 135005;
 Best Local Similarity 71.8%; Pred. No. 2.6e-28;
 Matches 260; Conservative 0; Mismatches 101; Indels 1; Gaps 1;

QY 3263 TTTTAAATTTTTTAAAAAATAGATGATGAGTGGTGGCTCATGCTGCTTAATCCCAACACTT 3322
 Db 26435 TTTTATAAAAAATGGTGACAGTGGCTGAGTGCTGCTCATGCTGCTTAATCCCAACACTT 26494
 QY 3323 TGGGAAGCCGGTTCGGGAGGATAGCTTTGAGTCCAGCAGTTTGAGACAGTCCAGTCCAGGCAACA 3382
 Db 26495 TGGGAGCTGAGCAGGGGATGCTTTGAGCCCATGAGTTTGAGACAGCTGCGCAACA 26554
 QY 3383 CAGCAAGACCCCATATCTAAAAAACAACAAAAATTAACCTGGGTATGGTTGTGC 3442
 Db 26555 GGGTAAAAATCCCATCTCAACAAAAAATACACACAAAAAATAGCTGGGCATGGTGTGT 26614
 QY 3443 TCACCTGTAGTCCAGCTACACAGGAGCTGAGGAGGAGGATCAGTTGAGCCCAAGAGG 3502
 Db 26615 ACATCTGTGTCCAGCTACTCAGGAGGCGGAGGAGGATCAGCTGAGCCCAAGAGG 26674
 QY 3503 TTGAGGCTGAGTCCAGCTGAGGAGGCTGCTACACT-CAGTCTGGGTGACAGTGCAGA 3561
 Db 26675 TCTAGACTGAGGAGCTGTGATTTGCCACTCCACTTCCAGTCTGGGCAACAGAGTGAGA 26734
 QY 3562 AGCTGTCTCAAAATATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 3621
 Db 26735 CCCTGTCTAAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 26794
 QY 3622 TT 3623
 Db 26795 AT 26796

RESULT 6
 US-10-380-124-10
 ; Sequence 10, Application US/10380124
 ; Publication No. US20040053874A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Isis Pharmaceuticals, Inc.
 ; APPLICANT: Brett P. Monia
 ; APPLICANT: Susan M. Freier
 ; TITLE OF INVENTION: ANTISENSE MODULATION OF CLUSTERIN EXPRESSION
 ; FILE REFERENCE: RTS-0156
 ; CURRENT APPLICATION NUMBER: US/10/380,124
 ; CURRENT FILING DATE: 2003-03-10
 ; NUMBER OF SEQ ID NOS: 90
 ; SEQ ID NO 10
 ; LENGTH: 8133
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-10-380-124-10

Query Match 3.3%; Score 187.6; DB 18; Length 8133;
 Best Local Similarity 74.6%; Pred. No. 9.3e-29;
 Matches 249; Conservative 0; Mismatches 84; Indels 1; Gaps 1;

QY 3264 TTTTAAATTTTTTAAAAAATAGATGATGAGTGGTGGCTCATGCTGCTTAATCCCAACACTT 3323
 Db 1057 TTTTAAAGAGGAAAAAATCTGAGCTGGGCATGGTGGCTCATGCTGCTGATGCCAGACTTT 1116
 QY 3324 GGGAGCCGGTTCGGGAGGATAGCTTGAGTCCAGCAGTTTGAGACAGTCCAGGCAACAC 3383
 Db 1117 GAGAGCCCAAGGCGGAGGATCATTTGAGCCAGGAGTTTGAGACAGCTGCGCAAGAT 1176
 QY 3384 AGCAAGACCCCATATCTAAAAAACAACAAAAATTAACCTGGGTATGGTTGTGCT 3443
 Db 1177 AGTGAGACCTGTCTCTACAAAATAATAATAATAATAATAATAATAATAATAATAATA 1236
 QY 3444 CACTGTAGTCCAGCTACACAGGAGCTGAGGAGGAGGATCACTTGAGCCCAAGGAGGT 3503
 Db 1237 CACCTGTAGTCTCAGCTACTCGGAGGCTGAGGCAAGAGAGTCACTTGAGCTGGAAGTT 1296
 QY 3504 TGAGGCTGCAGTGTATCCATGACCGCTGCTACACT-CAGTCTGGGTGACAGTCCAGAA 3562
 Db 1297 GGAGGCTGCAGTGCAGTGTATGATTCACCATTCAGCTTCCAGCTGGGCAACAGAGTGAGAC 1356
 QY 3563 GCTGTCTCAAAAAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 3596


```

; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: fastSEQ for Windows Version 4.0
; SEQ ID NO 115440
; LENGTH: 3287
; TYPE: DNA
; ORGANISM: Human
; US-10-027-632-115440

Query Match          3.3%; Score 185.4; DB 13; Length 3287;
Best Local Similarity 76.3%; Pred. No. 1.7e-28;
Matches 254; Conservative 0; Mismatches 76; Indels 3; Gaps 2;

QY 3265 TTAATAATTTTAAAAAATAGATGAGTGTGGTTCATCGCTGTATCCCAACACTTTG 3324
DB 843 TTTACTTTTTTAATAAGCGAGCTGGGTGCAGTGGCTCACCTGTATCCGACACTTTG 784

QY 3325 GGAAGCGGGTCGGGAGGATAGCTTGAGTCCAGCAGTTTGAGACCAGCTCAGGGCAACACA 3384
DB 783 GGAGGCCAAGTGGTGTGATTGTTGAGTTTCAAGAGTTTGCAGACCACCTGGGCAACATG 724

QY 3385 GCAAGACCCCATATCTAAAAAACAACAAAAAATAAATTAACCTGGGTATGGTTGTGCTC 3444
DB 723 GCGAAAAACCATCTCTACAAAA--AATACAAAAAAGATTAGCGGGTGTGGTAGCAGAC 666

QY 3445 ACCTGTAGTCCAAGCTACACAGGAGCTGAGGAGAGGATCACTTCAGGCCCAGGAGTT 3504
DB 665 ACCTGTAGTCCCAGCTACTTAGGAGGCTCAGGTGGGAGGATCACTGAGGCCCAGGAGTT 606

QY 3505 GAGCTCAGTGTATCCATGAACGGCTGTACACT-CAGTCTGGGTGACAGTGCAGAAG 3563
DB 605 GAAGCTCGATGAGCTGTGATCGTGCACCTGCACTCCAGACTGGGTGATAGTGCAGACA 546

QY 3564 CTGTCTCAAAAAATAAATAAATAAATAAATAAAC 3596
DB 545 CTATCTCAAAAAAATAAATAAATAAATAAATAAGAC 513

RESULT 11
US-10-027-632-115440/c
; Sequence 115440, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; POLYMERIZATION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/219,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: fastSEQ for Windows Version 4.0
; SEQ ID NO 115440
; LENGTH: 3287

```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 22, 2005, 14:47:40 ; Search time 16574 Seconds
(without alignments)
13035.646 Million cell updates/sec

Title: US-10-030-294-1

Perfect score: 5676

Sequence: 1 cagctgttcaggatgctg.....tttataataattgtctccgt 5676

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

- 1: gb_est1.*
- 2: gb_est2.*
- 3: gb_btc.*
- 4: gb_est3.*
- 5: gb_est4.*
- 6: gb_est5.*
- 7: gb_est6.*
- 8: gb_gss1.*
- 9: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	427.8	7.5	523	8	A0168656	HS 3165 B
2	293.2	5.2	638	6	BY723946	BY723946 BY723946
3	280.4	4.9	293	4	BG212218	BG212218 RST31802
C 4	267	4.7	288	4	BG219090	RST31802
C 5	262	4.6	277	4	BG212753	RST31802
6	184.6	3.3	567	4	B1497128	df132c05
C 7	181.6	3.2	664	8	A0275866	CITBI R1-
8	180.6	3.2	725	8	CC061176	MUCQ CH25
9	180	3.2	828	8	B2601705	WHABF1TR
10	179.8	3.2	733	8	AQ052879	RPC111-42
11	179.6	3.2	567	5	B0691256	AGENCOURT
12	179.6	3.2	810	8	A0780966	HS 3138 B
13	179	3.2	637	5	B0688837	AGENCOURT
C 14	178.8	3.2	440	1	A1362915	qy81c06.x
C 15	178.2	3.1	674	9	AG073447	Pan trogl
16	178	3.1	463	5	BQ883640	AGENCOURT
C 17	177.8	3.1	449	1	A1365624	ap20e10.x
18	177.6	3.1	680	8	A0750933	HS 5574 B
19	177.6	3.1	1025	5	BM907570	AGENCOURT
20	176	3.1	680	8	AQ046675	UI-E-CQ1-
C 21	176	3.1	706	4	BM701690	UI-E-CQ1-
22	175.4	3.1	330	1	AA226173	nc11908.8
23	175.4	3.1	626	5	BX952768	DKFZp781E
24	175.4	3.1	1710	3	CR590862	full-leng

C 25	175	3.1	917	5	BUI53359	AGENCOURT
26	174.2	3.1	657	8	AQ343025	RPC111-12
27	174.2	3.1	834	8	AQ749761	HS 5573 A
28	174	3.1	490	6	CA395712	cs68a08.y
C 29	173.6	3.1	360	1	AI167878	oz91b12.x
C 30	173.4	3.1	677	6	CA307465	UI-H-F11-
C 31	173.2	3.1	680	7	CR546471	DKFZp470B
32	173.2	3.1	5925	3	CR749594	Homo sapi
33	173.2	3.1	6069	3	BC036681	Homo sapi
C 34	172.8	3.0	532	8	CC061208	MUCQ CH25
35	172.4	3.0	589	2	BE562634	601336472
C 36	172.2	3.0	693	9	AG087187	Pan trogl
37	172	3.0	549	7	W07317	za94b10.r1
C 38	171.8	3.0	416	1	AA715814	nm25c01.s
39	171.8	3.0	689	7	CN345184	170005999
40	171.8	3.0	4735	3	HS0803160	Homo sapi
41	171.6	3.0	409	8	AQ413740	RPCI-11-2
42	171.6	3.0	550	7	CR545302	DKFZp459I
43	171.6	3.0	586	8	AQ038047	CIT-HSP-2
C 44	171.6	3.0	1105	1	AL543301	AL543301
45	171.4	3.0	729	4	BG536633	602566240

ALIGNMENTS

RESULT 1
LOCUS A0168656 523 bp DNA linear GSS 16-OCT-1998
DEFINITION HS_3165_B2_F08_T7 CIT Approved Human Genomic Sperm Library D Homo sapiens Genomic clone Plate=3165 Col=16 Row=L, genomic survey sequence.
ACCESSION A0168656
VERSION A0168656.1 GI:3566331
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 523)
AUTHORS Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T., and Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D., and Hood,L.
TITLE Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome
JOURNAL Proc.Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
MEDLINE 99380589
PUBMED 10449764
COMMENT Contact: Mahairas GG, Wallace JC, Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu
Sequence Tagged Connector
Plate: 3165 row: L column: 16
Class: BAC ends
High quality sequence stop: 523.
Location/Qualifiers
source
1. 523
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="plate=3165 Col=16 Row=L"
/sex="male"
/note="lib="CIT Approved Human Genomic Sperm Library D"
/note="Organ: sperm; Vector: pBelOBAcl1; BAC Clones in E-Coli DH10B"

Query Match 7.5%; Score 427.8; DB 8; Length 523;
Best Local Similarity 90.6%; Pred. No. 1.4e-59;

	Matches	453;	Conservative	0;	Mismatches	47;	Indels	0;	Gaps	0;
Qy	672	GATCTCATAGTCA	CAGGCTCC	GAGCTTAC	GGAAACAT	CAAGATCAT	CTGCTTTAAGC	731		
Db	12	GANTCNTAGTAC	GGATGAAT	TCGATCTAC	GGAAACAT	CGAATCAT	CTGCTNTAAGC	71		
Qy	732	CTCTTGTTGG	GAATCTGTTT	TCACCTGC	CCCTGTAC	CTATTGCTCT	TTTCCCTTGTTA	791		
Db	72	CTCTTGTTGG	CAATCTGTTT	TCACCTGC	CCCTGTAC	CTATTGCTCT	TTTCCCTTGTTA	131		
Qy	792	ACAGAACCTTT	ATTCTTCTG	AAAACTCT	CTGCTCAGT	CATGGTAGGG	CCATCAGTCCA	851		
Db	132	ACAGAACCTTT	ATTCTTCTG	AAAACTCT	CTGCTCAGT	CATGGTAGGG	CCATCAGTCCA	191		
Qy	852	CATGATCAGG	CCCTCTCTG	GCCAAA	CAATGCA	TCTTTCTTTT	TGGGAATTTGAATCTTAAAG	911		
Db	192	CATGATCAGG	CCCTCTCTG	GCCAAA	CAATGCA	TCTTTCTTTT	TGGGAATTTGAATCTTAAAG	251		
Qy	912	CTGAATAGCT	GGAAGTTCA	AAAAAGCT	GTGTTGAAT	CTGACTTAC	GCCCTACAGTGGCTTTGC	971		
Db	252	CTGAATAGCT	GGAAGTTCA	AAAAAGCT	GTGTTGAAT	CTGACTTAC	GCCCTACAGTGGCTTTGC	311		
Qy	972	AAAGTGACTG	TCCATTCCT	ATTCTTAA	GTCCTGAA	TTTATAAT	TATCTCTGGTTACAG	1031		
Db	312	AAAGTGACTG	TCCATTCCT	ATTCTTAA	GTCCTGAA	TTTATAAT	TATCTCTGGTTACAG	371		
Qy	1032	CCCTTTCTGA	GATGTGGT	TTTTTTTCC	AACTGTCCTT	TATAGTCT	GTGTAATTTTCATA	1091		
Db	372	CCCTTTCTGA	GATGTGGGA	TTTTTTTCC	AACTGTCCTT	TATAGTCT	GTGGAATTTATCATA	431		
Qy	1092	TTTCTTTTCA	TACATTTTC	ATGTTTTG	TTTGTGTTG	TGTTTTG	CTTTTGGCTTTAGGTA	1151		
Db	432	TTTCTTTTCA	TACATTTCA	TGCTGTTG	TTTGTGTTG	TGTTTTG	CCGTGATGNTGTCAGACTTAGTA	491		
Qy	1152	GGCAGATCA	GTTTCTG	TGTG	1171					
Db	492	GAAGAATG	ATCTATT	CTG	511					

RESULT 2	BY723946	638 bp	linear	EST 17-DEC-2002
LOCUS	BY723946			
DEFINITION	RIKEN full-length enriched, adult male hypothalamus Mus musculus cDNA clone A230061P11 5', mRNA sequence.			

ACCESSION VERSION KEYWORDS SOURCE ORGANISM	REFERENCE AUTHORS
BV723946 BX723946.1 GI-27137063 EST. Mus musculus (house mouse) Mus musculus Mammalia; Eutheria; Rodentia; Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus	1 (bases 1 to 638) Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S., Nikaido, I., Osato, N., Saito, R., Suzuki, H., Yamanaka, I., Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Kiyonobach, C., Gojobori, T., Baldarelli, R., Hill, D. P., Bult, C., Hume, D. A., Quackenbush, J., Schriml, L. M., Kanpin, A., Matsuda, H., Batalov, S., Belsel, K. W., Blake, J. A., Bradt, D., Brusci, V., Chothia, C., Corbani, L. E., Cousins, S., Dalla, E., Dragani, T. A., Fletcher, C. F., Forrest, A., Frazer, K. S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grummond, S., Gustincich, S., Hirokawa, N., Jackson, I. J., Jarvis, E. D., Kanai, A., Kawaji, H., Kawasawa, Y., Kedzierski, R. M., King, B. L., Konagaya, A., Kurochkin, I. V., Lee, Y., Lenhard, B., Lyons, F. A., Maglott, D. R., Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T., Numata, K., Okido, T., Pavan, W. J., Perte, G., Pesole, G., Petrovsky, N., Pillai, R., Pontius, J. U., Qi, D., Ramachandran, S., Ravasi, T., Reed, J. C., Reed, D. J., Reid, J., Ring, B. Z., Ringwald, M., Sandelin, A., Schneider, C., Sample, C. A., Setou, M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M. S., Teasdale, R. D., Tomita, M., Verrardo, R., Wagner, L., Wallested, C. C., Wang, Y., Watanabe, Y., Wells, C., Wilming, L. G., Wysshaw-Boris, A., Yanagisawa, M., Yang, I., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P.,

```
Location/Qualifiers
1. .293
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/cell_line="HT1080"
```

Query Match 5.2%; Score 293.2; DB 6; Length 638;
Best Local Similarity 77.6%; Pred. No. 1.1e-37;

`/clone_lib="Athersys RAGE Library"`
`/notes="See 'Creation of Genome-wide Protein Expression Libraries using Random Activation of Gene Expression', Nature Biotechnology, in press. Note that even though the cell type indicated is Hri1080, since a random activation method was used, these sequence tags are not necessarily expressed in Hri1080 under normal circumstances."`

```
Query Match      4.9%; Score 280.4; DB 4; Length 293;
Best Local Similarity 99.6%; Pred. No. 1.6e-35;
Matches 281; Conservative 0; Mismatches 1; Indels 0
```

Qy	3117	TCTTTTTCATAATTACTTATTTAAATTTGTGGCAATTAATTTCTGTTCTCTTAGCTT	3176
Db	12	TTTTTTTCATAATTACTTATTTAAATTTGTGGCAATTAATTTCTGTTCTCTTAGCTT	71

Qy	3177	AGTAACTTTAGGATTTTTAAATAACAACTATTTGAATCATGACATACGTTTAAATGATAT	3236
Db	72	AGTAACTTTAGGATTTTTAAATAACAACTATTTGAATCATGACATACGTTTAAATGATAT	131

Qy	3237	TATTTAAATACGTTAGGCTATAAAACCTTTTAAATTTTTTAAAAAATAGATGAGTGTGGT	3296
Dβ	132	TATTTAAATACGTTAGGCTATAAACCTTTTAAATTTTTTAAAAAATAGATGAGTGTGGT	191

Qy	3297	GGCTCATGCCTGTATAATCCCAACACATTTTGGGAAGCCCGGTTCGGAGGATAGCTTGAGTCCA	3356
Dh	192	GGCTCATGCCTGTATAATCCCAACATTTTGGGAAGCCCGGTTCGGAGGATAGCTTGAGTCCA	251

QY		3357	GCAGTTTGAGACCACTCAGGGCAACA	CAGCAAGACCCCATAT	3398
nB		252	GCAGTTTGAGCCAGTCAGGCGACAC	CAGCAGACCCCATAT	293

RESULT 4
BC319090/c

LOCUS	RG219090	EST38843	Athersys RAGE Library	Homo sapiens	cdNA, mRNA	linear	288 bp	EST 21-APR-2000
DEFINITION	RG219090							
ACCESSION	RG219090							
VERSION	RG219090.1							
KEYWORDS	EST.							
SOURCE	Homo sapiens (human)							

ORGANISM

1 (bases 1 to 288)

REFERENCE

AUTHORS

Harrington, J. J., Sherf, B., Rundlett, S., Jackson, P. D., Perry, R.,
Cain, S., Leventhal, C., Thornton, M., Ramachandran, R.,
Whittington, J., Lerner, L., Costanzo, D., McElligott, K., Booser, S.,
Mays, R., Smith, E., Veloso, N., Klika, A., Hess, J., Cochran, K., Lo, K.,
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

TITLE
Creation of genome-wide protein expression libraries using random
ORF libraries
Oriánbacher, J., Danzig, J. and Bucar, M.

JOURNAL
Nat. Biotechnol. 19 (5), 440-445 (2001)

MEDLINE
21227151

PUBMED
11329013

FEATURES
 source
 1..288
 Location/Qualifiers
 High quality sequence stop: 286.
 Email: scain@athersys.com
 Fax: 216 361 9596
 Tel: 216 431 9900
 Athersys, Inc.
 3201 Carnegie Ave, Cleveland, OH 44115, USA

	ORIGIN	Query Match	4.7%; Score 267; DB 4; Length 288;	
		Best Local Similarity 98.2%; Pred. No. 2.4e-33;		
		Matches 270; Conservative 0; Mismatches 5; Indels 0; Gaps 0;		
QY	3119	TTTTTGCATAATTACTTATTTAAATAATTTGTTGCCATTTAAATTCCTGTCCTCTTAGCTTAG	3178	
Db	287	TTTTTTTTTTTTTACTTAATTTAAATAATTTGTTGCCATTTAAATTCCTGTCCTCTTAGCTTAG	228	
QY	3179	TAACTTTAGGATTTTTTAAATAACAACTATTGAAATCATGACATACGTTTAAATGATATTA	3238	
Db	227	TAACTTTAGGATTTTTTAAATAACAACTATTGAAATCATGACATACGTTTAAATGATATTA	168	
QY	3239	TTTTAAATACGTTAGGCTATAAACCTTTTAAATTTTTTAAAAAANTAGATGAGTGTGGTGG	3298	
Db	167	TTTTAAATACGTTAGGCTATAAACCTTTTAAATTTTTTAAAAAANTAGATGAGTGTGGTGG	108	
QY	3299	CTCATGCTCTGTAATCCCAACACACTTTGGGAAGCCGGGTCCGGAGGATAGCTTTGAGTCCAGC	3358	
Db	107	CTCATGCTCTGTAATCCCAACACACTTTGGGAAGCCGGGTCCGGAGGATAGCTTTGAGTCCAGC	48	
QY	3359	AGTTTGAGACCAAGTCAGGGCAACACAGAACCC	3393	
Db	47	AGTTTGAGACCAAGTCAGGGCAACACAGAACCC	13	

ORIGIN		Query Match	4.6%; Score 262; DB 4; Length 277;	
		Best Local Similarity 98.1%; Pred. No. 1.6e-32;		
		Matches 265; Conservative 0; Mismatches 5; Indels 0; Gaps 0;		
Qy	3117	TCATTTTCATAAATACCTATTATAAAATTTGTGGCCATTTAAATTTCTGTCTCTTAGCTT	3176	
Db	270	TTTTTTTCATAAAATACCTATTATAAAATTTGTGGCCATTTAAATTTCTGTCTCTTAGCTT	211	
Qy	3177	AGTAACTTTTAGGATTTTAAATAACAACACTATTGAAATCATGACATACGTTTAAATGATAT	3236	
Db	210	AGTAACTTTTAGGATTTTAAATAACAACACTATTGAAATCATGACATACGTTTAAATGATAT	151	
Qy	3237	TATTTAAATACGTTTAGGCTATAAACCTTTTAAATTTTTTAAAAAAATAGATGAGTGTGGT	3296	
Db	150	TATTTAAATACGTTTAGGCTATAAACCTTTTAAATTTTTTAAAAAAATAGATGAGTGTGGT	91	
Qy	3297	GGCTCATGCGTGTAAATCCCAACACTTTTGGGAAGCCGGGTGCGGAGGATAGCTTTGAGTCCA	3356	
Db	90	GGCTCATGCGTGTAAATCCCAACACTTTTGGGAAGCCGGGTGCGGAGGATAGCTTTGAGTCCA	31	
Qy	3357	GCAGTTTGGACACGATCAGGGCAACACAGC	3386	
Db	30	GCAGTTTGGACACGATCAGGGCAACACAGC	1	

RESULT 5	BG212753/c	777 bp	mRNA	linear	EST 21-APR-2001
LOCUS	RG212753.1				
DEFINITION	Athersys RAGE Library Homo sapiens				
ACCESSION	BG212753				
VERSION	BG212753.1	GI:13734440			
KEYWORDS	EST.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
	Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.				
REFERENCE	1 (bases 1 to 277)				
AUTHORS	Harrington, J.J., Sherf, B., Rundlett, S., Jackson, P.D., Perry, R.,				
	Chin, S., Leventhal, C., Thornton, M., Ramchandran, R.,				
	Whittington, J., Lerner, L., Costanzo, D., McElligott, K., Booser, S.,				
	Mays, R., Smith, E., Veloso, N., Klika, A., Hess, J., Cothren, K., Lo, K.,				
	Offenbacher, J., Danzig, J., and Ducar, M.				
TITLE	Creation of genome-wide protein expression libraries using random				
	activation of gene expression				
JOURNAL	Nat. Biotechnol. 19 (5), 440-445 (2001)				
MEDLINE	21227151				
PUBMED	11329013				
COMMENT	Contact: Scott J. Cain				
	Athersys, Inc.				
	3201 Carnegie Ave, Cleveland, OH 44115, USA				
	Tel: 216 431 9900				
	Fax: 216 361 9596				
	Email: scain@athersys.com				
FEATURES	High quality sequence stop: 227.				
source	Location/Qualifiers				
	1..277				
	/organism="Homo sapiens"				
	/mol_type="mRNA"				
	/db_xref="taxon:9606"				
	/cell_line="H1080"				
	/clone_lib="Athersys RAGE Library"				
	/note="See 'Creation of Genome-wide Protein Expression				

Class: SAC ends.
Location/Qualifiers
1. .664

ORIGIN

REFERENCE	1. Luessli, C. & Z. Dore, C. Villeneuve, A., Lepage, P.,
AUTHORS	Minhas, R., Zhang, X., McKee, K., Ophoff, R.A., Fairbanks, L.A., Freimer, N.B.,
	Forgetta, V., Ervin, F.R., Palmour, R.M., Hudson, T.J. and Dewar, K.
TITLE	UCLA/MUGO/St-Kitts Vervet Monkey Mapping Project
JOURNAL	Unpublished (2003)
COMMENT	Other GSSs: MUGO_CH252P002Q4SP6_H05_CD315_036 Contact: Dewar K

McGill University and Genome Quebec Innovation Centre
 McGill University
 740 Dr. Penfield Room 7214, Montreal, QC, Canada, H3H 1A4
 Tel: 514 398 3311 x00089
 Fax: 514 398 1795
 Email: ken.dewar@mcgill.ca
 Seq primer: T7 : TAATACGACTCATATAGGG
 Class: BAC ends.

FEATURES

source Location/Qualifiers
 1. 725
 /organism="Cercopithecus aethiops"
 /mol_type="genomic DNA"
 /db_xref="taxon:9534"
 /clone="CH252-2B10"
 /sex="male"
 /cell_type="White blood cell"
 /dev_stage="Adult"
 /clone_lib="CHORI-252 Vervet Monkey Library"
 /note="Vector: pFARAC2.1; Site 1: EcoRI; Site 2: EcoRI;
 Constructed by Michael Nefedov in Pieter de Jong's
 laboratory at BACPAC Resources, Children's Hospital in
 Oakland Research Institute."

ORIGIN

Query Match 3.2%; Score 180.6; DB 8; Length 725;
 Best Local Similarity 60.1%; Pred. No. 2.3e-19;
 Matches 335; Conservative 0; Mismatches 219; Indels 3; Gaps 2;

Qy 3078 AACTCTCGCCATTATTTCCATATGCAATAGGACCAATCTTTTCATATTAATCTATT 3137
 Db 21 AACACCCTAACTTTATTGGCAAGCTTGTATATTTCCAAAGTATATATTTTGTGTT 80
 Qy 3138 TAAATTTGTGCCATTATTTCTGTCCTCTTAGCTTAGTAATTTAGATTTTTAAA 3197
 Db 81 TAAACATGCTATTAAATTTTACAGATTAATTTTCTTTTGTGCTGAGGTGAAGAACA 140
 Qy 3198 TAACAACTATTGAATCATGACATACGTTAAATGATATTAATTAATAGTTAGGCTAT 3257
 Db 141 TGTCTGTATGTTATGCTTTGTAATTTATTAATCTTTTGTGACCTGGTGTAATA 200
 Qy 3258 AAACCTTTTAAATTTTAAATAATAGATGAGTGTGGTGCATGCTGTATATCCCAA 3317
 Db 201 ACAATGAACAATAATTGTATGATACTGCTGGGCACGATAGCTCATGCTGTATCCAG 260
 Qy 3318 CACTTTGGGAAGCGGTGCGGAGGATAGCTTGAGTCAGAGTTTGAGACCAGTCAAGG 3377
 Db 261 CAGTTTGGGAGGCCAAGGTGGTGGATCACTTGAGATCAGGAGTTCAGACCAGCCTGGC 320
 Qy 3378 CAACACAGCAAGACCCATATCTATAAAAAAACAACAAACAAATTTACCTGGGTATGGT 3437
 Db 321 CAACATGGCAAACTCATCTCTACTATAAAAAAACAACAAATAA--TAGCTGGGTGGT 378
 Qy 3438 TGTGCTCACCTGTAGTCCAGTACACAGGAGCTGAGGCAAGGATCACTTGAGCCCA 3497
 Db 379 TGCACACACCTGTAATCCAGATCTTTGGGAGCTGAGGCGAGGAGATCACTTAAATCTA 438
 Qy 3498 GGAGTTGAGGCTGAGTCCATGAACGGCTGTCTACT--CAGTCTGGGTGACAGTG 3556
 Db 439 GGAGGAGAGGTTACAATAGACCAAGATTGCACCACCTCCAGCTTGGTGACAGAG 498
 Qy 3557 CAAGAAGCTGTCTCAAAAAATAATAATAATAATAATAATAATAATAATAATAATAATA 3616
 Db 499 AGAGACACTGTCTCAAAAAAATAATAATAATAATAATAATAATAATAATAATAATA 558
 Qy 3617 TTAATTTTAAAAACAC 3633
 Db 559 TACAAATTTATGATAC 575

RESULT 9

BZ601705 828 bp DNA linear GSS 08-JUN-2003
 LOCUS

DEFINITION

WHABF61TR Human MCF7 breast cancer cell line library (MCF7_1) Homo sapiens genomic clone MCF7_1-8K2, genomic survey sequence.

ACCESSION BZ601705
 VERSION BZ601705.1 GI:31510167
 KEYWORDS GSS.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 828)
 AUTHORS Volik,S., Zhao,S., Chin,K., Brebner,J.H., Herndon,D.R., Tao,Q., Kowbel,D., Huang,G., Lapuk,A., Kuo,W.-L., Magrane,G., de Jong,P., Gray,J.W. and Collins,C.

TITLE End-sequence profiling: Sequence-based analysis of aberrant genomes
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 100 (13), 7696-7701 (2003)
 MEDLINE 22709111
 PUBMED 12788976

COMMENT Contact: Volik SV
 Colin Collins' lab
 UCSF Comprehensive Cancer Center
 UCSF Box 0808, San Francisco, CA 94143-0808, USA
 Tel: 415 502 7066
 Fax: 415 502 5665
 Email: svolik@cc.ucsf.edu
 This clone is available from Amplicon Express
 http://www.genomex.com
 Class: BAC ends.

Location/Qualifiers
 1. 828
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /clone="MCF7_1-8K2"
 /sex="female"
 /clone_lib="Human MCF7 breast cancer cell line library (MCF7_1)"
 /note="Vector: pECBAC1; Site 1: HindIII; This library was constructed from MCF7 breast cancer cell line by Amplicon Express (http://www.genomex.com) using their standard procedure."

Query Match 3.2%; Score 180; DB 8; Length 828;
 Best Local Similarity 67.7%; Pred. No. 2.9e-19;
 Matches 283; Conservative 0; Mismatches 130; Indels 5; Gaps 2;

Qy 3295 GTGGCTCATGCTGTATATCCAAACACTTTGGGAAGCGGTCGGGAGGATAGCTTAGTC 3354
 Db 401 GTGGTTCATACCTATAATCCCAACACTTTGGGAGACTGAGGCTGGAAGATCACTTGAGAC 460

Qy 3355 CACGAGTTTGAGACCTAGGCGCAACAGCAAGACCCCATATCTAAAAAACAACAA 3414
 Db 461 CAGGAGTTTGAGACCTAGGCGCAACAGCTAGTACCCCATCT--ACATAAAAAA 516

Qy 3415 AAACAAAAATTACCTGGGTATGTTGTCTCACCTGTAGTCCAAAGCTACACAGGAAGCTGA 3474
 Db 517 TAAATATTAGTAGGATAGTATGATATGACCTGTAGTACTAGTACTTTGGGAGGCTGA 576

Qy 3475 GGCAAGGATCACTTGAGCCCAAGAGGTTGAGGCTGAGGATCCATGAACCGGCTGCT 3534
 Db 577 GGCAAGGATCGCTTTGAGCCCAAGAGGTTGAGGCTGAGGCTGAGTGAATGTGCCACT 636

Qy 3535 ACACCTC-AGTCTGGGTGAGCTGCAAGAGCTCTCTCAAAATAATAATAATAATAATAA 3593
 Db 637 GCACCTCTAGGCTTGGAGACAGACAGATCTTCTCTCTCTCTCTCTCTCTCTCTCTCT 696

Qy 3594 AACCTTTAAAAAACAACAAATTAATTAATTTAAAAAACAACACACACTAGAGTGTTCG 3653
 Db 697 ATTTTAAAAAACAAGAGTCTTTGAACCTCTTGAGACCAAGACTATAATTTTAATCTAA 756

Qy 3654 AAATGATTAATTTGGGAGTCTATATCCCTGGAGAGTAAATTTAAATAATTTAGAGAGT 3711
 Db 757 AATCT 814

ORIGIN

RESULT 10
 A0052879 733 bp DNA linear GSS 20-APR-1999
 LOCUS
 DEFINITION
 A0052879
 A0052879.1 GI:3349834
 GSS.
 Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 733)
 Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K.,
 Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., de Jong,P. and
 Venter,J.C.
 Use of BAC End Sequences for Sequence-Ready Map Building (1998)
 Unpublished (1998)
 Contact: Mark Adams
 Department of Eukaryotic Genomics
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA
 Tel: 301 838 0200
 Fax: 301 838 0208
 Email: mdadams@tigr.org
 Clones are derived from the human BAC library RPCI-11. For BAC
 library availability, please contact Pieter de Jong
 (pieter@dejong.med.buffalo.edu). Clones may be purchased from
 BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
 Research Genetics (info@resgen.com). BAC end search page:
 http://www.tigr.org/tldb/hungen/bac_end_search/bac_end_search.html
 Class: BAC ends.

FEATURES
 source
 1. 733
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="GDB:7516102"
 /db_xref="taxon:9606"
 /clone="RPCI-11-42023"
 /sex="Male"
 /cell_type="Lymphocytes"
 /clone_lib="RPCI-11"
 /note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
 RPCI11 Human Male BAC Library"

ORIGIN
 Query Match 3.2%; Score 179.8; DB 8; Length 733;
 Best Local Similarity 65.9%; Pred. No. 3.2e-19;
 Matches 276; Conservative 0; Mismatches 142; Indels 1; Gaps 1;
 QY 3196 AATAACAACACTATTGAAATCATGACATAGCTTTAAATGATATTTAAATACGTTAGGCT 3255
 DB 267 ATTAAGAGATCATTTAATCCAGGCCATTTATTTCAATGAGAGAGCTAAACGGCAGAAAG 326
 QY 3256 ATAAACCTTTTAAATTTTAAATAAATAGATGAGTGTGGTGCATGCTGTAAATCCC 3315
 DB 327 TTAATATGACTTGTTCATGTTTCATATCATTTGGCTGGGCATGGTGGCTCACACCTGTAATTC 386
 QY 3316 AACACTTTGGGAGCCGGGTGGGAGGATAGCTTCAGTCCAGCAGTTTGGAGACCAGTCAG 3375
 DB 387 AGCACTTTGGGAGGCCCGAGGTGGAGGAAACATTGAGCCCGAGGATTCGAGATCAGCCTG 446
 QY 3376 GGCACACAGCAGAACCCCATATCTTAAATAAACAACAAACAAATTTACCTGGGTATG 3435
 DB 447 GNCACATGTTGAACCCCATCTCTTAAATAAATAAATAAATAAATTTAGCAGGCTTT 506
 QY 3436 GTTGTGCTACCTGTAGTCCAAAGTACACAGGAAGCTGAGGAGGATCATTGAGCC 3495
 DB 507 GCAGCACATGCTGTGGTCCCAGCTACTTGGAGGCTGAAGTGGGAGGATGCTTGGGCC 566
 QY 3496 CAGGAGGTGAGGCTGCAGTGATCCATGACCGGCTGCTACACT-CAGTCTGGGTGACAG 3554

DB 567 CAGGAGGTCAAGGCTACAGTGAACCTGATTGTACCACTACATCCAGCTGGGTGACAG 626
 QY 3555 TGCAGAAGCTGTCTCAAAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 3613
 DB 627 AGTGAACCTGTCTCTCAAAACAACAACAACAACAACAACAACAACAACAACAACAATA 685

RESULT 11
 BO691256 567 bp mRNA linear EST 15-JUL-2002
 LOCUS
 DEFINITION
 AGENCOURT_8176628 NIH_MGC_110 Homo sapiens cDNA clone IMAGE:6252453
 5', mRNA sequence.
 BO691256
 BO691256.1 GI:21816572
 EST.
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 567)
 NIH-MGC http://mgi.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LICM2398 row: 1 column: 22
 High quality sequence stop: 542.

FEATURES
 Location/Qualifiers
 1. 567
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6252453"
 /tissue_type="ductal carcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH MGC 110"
 /note="Organ: pancreas; Vector: pOTB7; Site 1: XhoI;
 Site 2: EcoRI, cDNA made by oligo-dT priming.
 Directionally cloned into EcoRI/XhoI sites using the
 following 5' adaptor: GGCACGAG(G). Library constructed by
 Ling Hong in the laboratory of Gerald M. Rubin (University
 of California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies).
 Note: this is a NIH_MGC Library."

ORIGIN
 Query Match 3.2%; Score 179.6; DB 5; Length 567;
 Best Local Similarity 73.0%; Pred. No. 3.6e-19;
 Matches 257; Conservative 0; Mismatches 88; Indels 7; Gaps 2;
 QY 3285 GATCAGTGTGTGGCTCATGCTGTATTCCTTAATCCAACTTTGGGAAGCCGGTGGGAGAT 3344
 DB 207 GCTGGGTACATGACTCAGCATCTGTAGTCCAGCAGCTTTGGGAGCCGAGCGGAGAT 266
 QY 3345 AGCTTGAGTCCAGCAGTTTGAGACCACTCAGGCGCAACACAGACAGCCCATATCTAAA 3404
 DB 267 CACTTGAGCCAGAGTTTGAGACCACTGCGCATGTGTGTGCTGTAGTCCAGCTGCTC 322
 QY 3405 AAACAAACAAACAAACAAATACCTGGGTATGTTGTGCTCACCTGTAGTCCAAGTACAC 3464
 DB 323 --ACCAAAATATGAATTTAGCTGGGCATGTGTGTGCTGTAGTCCAGCTGCTC 380
 QY 3465 AGGAAGCTGAGCGCAGAGAGATCATCTTGAGCCCAAGAGGTTGAGGCTGAGTATCCATGA 3524
 DB 381 AGGAGGCTGAGCGCGGAGGATGGCTTGAGCCCAAGAGGTTGAGGCTGAGGCTGAGCCATGA 440

Qy	3441	GCTTACCTGTAGTCCAAAGCTACACAGAAAGCTAGGACAGAAAGATCACTTGAGCCCAAGGA	3500
Db	440	ACATGCTGTGTAGTCCCAAGCTACTCATGAGCTGAGGACAGAGAATCACTTTGAACCCCAAG	499
Qy	3501	GGTTGAGGCTGCAGTGATCCATCAAGACCGCTGCTACACTCAGTCTGGGTGACAGTGCAG	3560
Db	500	GGCGGAGGTGTGAGTGGCGCGAGATTGTGCCACTGCACCTCCAGCTGGCGACAGAGCAAG	559
Qy	3561	AAAGCTGCTCAAAAAATAATAATAATAAAAAATAACCTTTTAAAAAACAATAATTA	3620
Db	560	ATTCTGTCTCAAAAAAATAATAATAATAATAATAATAATAATAATAATAATAATAATA	619
Qy	3621	ATTTTA 3626	
Db	620	CATATA 625	
RESULT 13			
BQ688837			
LOCUS	BQ688837		
DEFINITION	AGENCOURT_B344575 NIH_MGC_110 Homo sapiens cDNA clone IMAGE:6247825 5', mRNA sequence.		
ACCESSION	BQ688837		
VERSION	BQ688837.1		
KEYWORDS	GI:21814153		
SOURCE	EST.		
ORGANISM	Homo sapiens (human)		
REFERENCE	Homo sapiens		
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
TITLE	1 (bases 1 to 637)		
JOURNAL	NIH-MGC http://mgs.nci.nih.gov/		
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999) Contact: Robert Strausberg, Ph.D. Email: cgapbs-r@mail.nih.gov Tissue Procurement: ATCC cDNA Library Preparation: Rubin Laboratory cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov plate: LLCM2386 row: 1 column: 02 High quality sequence stop: 636. Location/Qualifiers 1. .637 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="IMAGE:6247825" /tissue_types="ductal carcinoma, cell line" /lab_host="DH10B (phage-resistant)" /clone_lib="NIH_MGC_110" /notes="EcoRI; pancreas; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dt priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."		
FEATURES			
source			
ORIGIN			
Query Match 3.2%; Score 179; DB 5; Length 637;			
Best Local Similarity 75.4%; Pred. No. 4.4e-19;			
Matches 251; Conservative 0; Mismatches 75; Indels 7; Gaps 2;			
Qy	3285	GATGAGTGTGGTTCATCGCTGTATCCCAACACTTTGGGAAGCGCGGTTCGGGAGGAT	3344
Db	311	GCTGGGTACAGTAACATCAATCTGTAGTCCACACACTTTGGGAGGCCGAGCGGGGCAGAT	370
Qy	3345	AGCTTGAGTCCAGCAGTTTGAGACCACTGAGGCAACACAGCAAGACCCCATATCTTAA	3404

Db 371 CACTTGAGCCAGGAGTTTGAGACCCAGCTGGGCAACACAGCGAAACCGTTTCTCT 426
QY 3405 AACAAACAAAACAAATACCTGGGTATGGTGTGCTCACCTGTAGTCCAAAGCTACAC 3464
Db 427 --ACCAAAATATGAATTTAGCTGGGCATGGTGTGCTGTAGTCCAGCTGCTC 484
QY 3465 AGGAAGCTGAGGAGAGGATCACTTGTAGCCCGAGGAGTTGAGCTGAGCTGATCATGA 3524
Db 485 AGGAGGCTGAGGCGGAGGAGTGGCTTGTAGCCCGAGGAGTTGAGCTGAGCTGAGCCATGA 544
QY 3525 ACGGCTGTCTACA-CTGAGTCTGGTGCACAGTGCAGAGAGCTGTCTCAAAATAATAAT 3583
Db 545 TTGGCCCACTGCACCCAGCTGGGCAACAGAGCAAGCCCTGTCTCAAAATAATATAT 604
QY 3584 AATATAAATAAATCTTTTAAATAAACAATAATTA 3616
Db 605 AATAAATAAAGTTTAAATAAATAAATAAATAA 637

RESULT 14
AI362915/c
LOCUS AI362915 440 bp mRNA linear EST 06-JAN-1999
DEFINITION QY81C06.x1 NCI CGAP Brn25 Homo sapiens cDNA clone IMAGE:2018410 3'
similar to contains Alu repetitive element;; mRNA sequence.
ACCESSION AI362915
VERSION AI362915.1 GI:4114536
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 440).
NCI/NINDS-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute / National Institute of Neurological
Disorders and Stroke, Brain Tumor Genome Anatomy Project
(CGAP/BTGP), Tumor Gene Index
Unpublished (1998)
AUTHORS Contact: Robert Straubeberg, Ph.D.
COMMENT Email: cgapbs-remail.nih.gov
Tissue Procurement: David N. Louis, M.D., Myrna R. Rosenfeld M.D.,
Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Seq primer: -40UP from Gibco.
FEATURES
source
1. .440
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2018410"
/tissue_type="anaplastic oligodendroglioma"
/lab_host="DH10B"
/clone_lib="NCI-CGAP Brn25"
/note="Organ: brain; Vector: p7T3D-Pac (Pharmacia) with a
modified polylinker; Site 1: Not 1; Site 2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer [5'
TGTTACCAATCTGAAGGGGAGCGCGCATAGTTTTTTTTTTTTTTTTTTT
T 3']; double-stranded cDNA was ligated to Eco RI
adaptors (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of the modified p7T3 vector.
Library is normalized, and was constructed by Bento
Soares and M. Fatima Bonaldo."

ORIGIN

Query Match 3.2%; Score 178.8; DB 1; Length 440;
Best Local Similarity 78.4%; Pred. No. 5e-19;
Matches 240; Conservative 0; Mismatches 62; Indels 4; Gaps 2;

QY 3290 GTGTGTGGCTATGCTGTATATCCAAACACTTTGGGAAGCCGGTCCGGAGGATAGCTT 3349
Db 304 GTGTGTGGCTATATCTGTATATCCAGCACTTTGGGAAGCCGAGATGGAGATCGCTT 245
QY 3350 GAGTCCAGAGTTTGTAGACAGCTCAGGCAACACAGAGCCCATATCTAAAAAACA 3409
Db 244 GAGCCAGAGTTTCAAGACAGCTCTGGGCAACATAGCAGACTCTCTCTACAAA--A 188
QY 3410 AAACAAAAACAAAATTAACCTGGGTATGTTGTCTCACCTGTAGTCCAAGTACACAGGAA 3469
Db 187 AATACAAAAAATTAGCAGGCTTGGTGGTGGTACCTTGTGTCCAGCTACTCTGGAG 128
QY 3470 GCTGAGCGAGAGGATCACTTGTAGCCCGAGGAGTTGAGCTGAGTCCATGACGCG 3529
Db 127 GCTGAGGTGGGAGATCACTTGTAGCCCGAGGAAATTTGAGGCTGCAATTCAGTATGATGCA 68
QY 3530 CTGCTACACT-CAGTCTGGGTGACAGTGCAGAGAGCTGTCTCAAAATAATAATAAATA 3588
Db 67 CAATTGCACCTCTGCTGGGAGACAGATGAGACTCTGTCTCAAAAAAAGAAAA 8
QY 3589 AAAATA 3594
Db 7 AGAAAA 2

RESULT 15
AG073447/c
LOCUS AG073447 674 bp DNA linear GSS 03-NOV-2001
DEFINITION Pan troglodytes DNA, clone: PTB-065D24.F, genomic survey sequence.
ACCESSION AG073447
VERSION AG073447.1 GI:16625249
KEYWORDS GSS.
SOURCE Pan troglodytes (chimpanzee)
ORGANISM Pan troglodytes
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
1
AUTHORS Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,
Totoki, Y., Watanabe, H. and Sakaki, Y.
COMMENT BAC end sequences of Library PTB
2 (bases 1 to 674)
Unpublished
Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,
Totoki, Y., Watanabe, H. and Sakaki, Y.
DIRECT SUBMISSION
Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suehiro-chou, Teurumi-Ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail: chimpses@gsc.riken.go.jp, URL: <http://hgp.gsc.riken.go.jp/>,
Tel: 81-45-503-9111, Fax: 81-45-503-9170)
Clones are derived from the chimpanzee BAC library PTB This BAC end
was generated during the R&D process and may have higher chance of
clone tracking errors.
PRIMERS
Sequencing: -21M13
LIBRARY
Vector : pKS145
R.Site 1 : SacI
R.Site 2 : SacI.
FEATURES
source
1. .674
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
/clone="PTB-065D24.F"
/sex="male"
/cell_type="lymphoblast"
/clone_lib="PTB Chimpanzee Male BAC Library"

ORIGIN

Query Match 3.1%; Score 178.2; DB 9; Length 674;
Best Local Similarity 75.9%; Pred. No. 5.8e-19;
Matches 233; Conservative 0; Mismatches 73; Indels 1; Gaps 1;

Qy	3289	AGTGTGGTGGCTCATGCTGTATCCCAACACTTTTGGGAAGCCGGGTCCGGAGGATAGCT	3348
Db	396	AGTGCAGTGGCTCATGCTGTATCCCAACATTTTGGTAGCCNAGCCGGAGGATCACT	337
Qy	3349	TGAGTCCAGCAGTTTGAGACCAGCTCAGGGCAACACAGCAAGCCCATATCTTAAAAAAC	3408
Db	336	TGAGTCCAGCAGTTTGAGACCAGCTCAGGGCAATACAGAAAGATCCTATTCTACCAAAA	278
Qy	3409	AAAAACAAAAAATTAACCTGGGTATGGTTGTGCTCACCTGTAGTCCAAGCTACACAGGA	3468
Db	277	ATAAAAAAATAAATAATGAGCCAGGAGTCTGTGCAATTCCTGTAGTCCAGCTACTCAGGA	218
Qy	3469	AGCTGAGGCAGAAAGGATCACTTGAGCCCAAGGAGTTTGAGGCTGCAGTGTCCATGAACGC	3528
Db	217	GTCGAGGTGGGAGGATCACTTGAGCCCGGAGTGTGAGGCTACACTAAGTGGTGATCAT	158
Qy	3529	GCTGCTACACTCAGTCTGGGTGACAGTGCAGAGAGCTGTCTCAAAAAATAAATAAATA	3588
Db	157	GCCACTGCACCTCAATCTGGGTGATAGAGTGAGACCCCTGTCTCAAAAAAATTTAAAAAGA	98
Qy	3589	AAATAA	3595
Db	97	ACTATGA	91

Search completed: June 23, 2005, 12:27:16
Job time : 16560 secs